Safety Science 152 (2022) 105778

Contents lists available at ScienceDirect

Safety Science

journal homepage: www.elsevier.com/locate/safety

A systematic review of human-AI interaction in autonomous ship systems

Erik Veitch *, Ole Andreas Alsos

NTNU, Department of Design, Kolbjørn Hejes Vei 2b, 7491 Trondheim, Norway

ARTICLE INFO

Keywords:
Automation
Artificial Intelligence
Work
Safety
Marine Navigation
Human-Computer Interaction
Safety management
Resilience Engineering
Interaction Design
Maritime Autonomous Surface Ships
STPA
Bayesian Networks

ABSTRACT

Automation is increasing in shipping. Advancements in Artificial Intelligence (AI) applications like collision avoidance and computer vision have the potential to augment or take over the roles of ship navigators. However, implementation of AI technologies may also jeopardize safety if done in a way that reduces human control. In this systematic review, we included 42 studies about human supervision and control of autonomous ships. We addressed three research questions (a) how is human control currently being adopted in autonomous ship systems? (b) what methods, approaches, and theories are being used to address safety concerns and design challenges? and (c) what research gaps, regulatory obstacles, and technical shortcomings represent the most significant barriers to their implementation? We found that (1) human operators have an active role in ensuring autonomous ship safety above and beyond a backup role, (2) System-Theoretic Process Analysis and Bayesian Networks are the most common risk assessment tools in risk-based design, and (3) the new role of shore control center operators will require new competencies and training. The field of autonomous ship research is growing quickly. New risks are emerging from increasing interaction with AI systems in safety–critical systems, underscoring new research questions. Effective human-AI interaction design is predicated on increased cross-disciplinary efforts, requiring reconciling productivity with safety (resilience), technical limitations with human abilities and expectations (interaction design), and machine task autonomy with human supervisory control (safety management).

1. Introduction

Artificial Intelligence (AI) and automation have the potential to improve safety in complex transportation systems. Unlocking that potential, though, appears to depend on a seamless integration of human and machine control, combined with well-executed strategies for managing risks in a constantly changing environment. Maritime Autonomous Surface Ships (MASSs) have in recent years emerged as a new application of vehicle automation, in turn presenting new challenges and a productive research community. In the backdrop of rapid technology development, research contributions about MASSs have come from disparate fields, including risk and safety science, human factors, policy, and engineering. Currently, there is no comprehensive review available to serve as a guidance across the multiple disciplines, despite what unites them all as a common goal: the vision of safe, effective MASS operations. The objective in this study is to systematically review the different disciplines contributing to MASS research. By identifying research themes and directions of current work, our motivation is to promote cross-pollination of research ideas and multi-disciplinary efforts as the field converges towards its common goal. Our aim in analyzing the research state-of-the-art is to contribute to a better understanding of current design frameworks for human-AI system integration. We also aim to synthesize the methods risk scientists are using to respond to the need to predict behavior of such systems in the face of considerable uncertainly and complexity.

Despite continually improving safety records, shipping is considered a dangerous industry with a high rate of fatal injuries and high consequences of maritime disasters (Hansen et al., 2002; Hetherington et al., 2006). The International Maritime Organization (IMO), the United Nations agency regulating international maritime safety, calls it "perhaps the most international of all the world's great industries - and one of the most dangerous" (IMO, 2019). Investigations of underlying causes for marine accidents tend to point to "human error" as the single greatest contributor, by some estimates involved in 75–96% of all accidents (Rothblum, 2000). Proponents of automation, harkening its precision and untiring capabilities, have long promised an end to human errors. And yet the statistics appear unchanged: one recent safety report stated that the same "75% to 96%" of marine incidents in 2020 were at least

* Corresponding author.
E-mail addresses: erik.a.veitch@ntnu.no (E. Veitch), oleanda@ntnu.no (O. Andreas Alsos).

https://doi.org/10.1016/j.ssci.2022.105778
Received 1 May 2021; Received in revised form 29 March 2022; Accepted 4 April 2022
Available online 12 April 2022

partly due to human error (Allianz, 2020). The tendency for automation applications to introduce new modes of error, even while effectively addressing known ones, has been well documented, including in the case of ship navigation (Lützhöft and Dekker, 2002). The risk picture emerging in the wake of MASS development remains unclear, hanging in the balance as new risks latent in human-AI interactions appear only partly offset by reduced instances of human error.

While it is difficult to measure the direct effect of automation on overall safety, comparisons of yearly safety reports show that safety at sea is generally improving (Allianz, 2020). Automation in shipping has incrementally increased since the transition of sails to engines and manual rudder control to powered steering. Today, "autopilot" system controls analogous to those on cars are commonplace on ships, as is Dynamic Positioning (DP) control in vessels like ferries, offshore supply vessels, survey ships, cable layers, and drill ships that require precise maneuvering. While automation and safety seem to advance in parallel, it is unclear to what extent the two are linked, and if so by what mechanisms.

At the start of the 2010s, a shift was observed towards digital transformation of the maritime industry, captured by the moniker "Maritime 4.0" and referring especially to automated integration of real-time data into decision making (Sullivan et al., 2020). The first major project about autonomous ships was completed in the years 2012–2015, called the MUNIN project (Burmeister et al., 2014). In 2017, two Norwegian companies Yara and Kongsberg announced plans to develop the Yara Birkeland, an autonomous cargo carrier servicing three ports in Southern Norway (Yara, 2017). Today, the project is still a forerunner, with plans for "fully autonomous operation" slated for 2022 (Kongsberg, 2020a). Recent innovations have demonstrated so-called "auto-crossing" and "auto-docking" in several ferry applications (Kongsberg, 2020b; Rolls-Royce, 2018). Collectively, these technologies effectively handle entire crossings between terminals automatically, relying on bridge crew to intervene to avoid possible collisions.

Today, AI is heralding the next level of automation in transportation applications, and it is the most recent harbinger of improved system safety addressing the ubiquitous "human error" problem. Deep learning has been demonstrated in applications ranging from obstacle detection (Chen et al., 2020), obstacle avoidance (Cheng and Zhang, 2018), and mooring line fault detection (Chung et al., 2020). As a research field, AI has been characterized by periods of rapid advancement and abundant research funding (so-called "AI Springs") followed by periods of slower progress ("AI Winters"). At the start of the 2010s, new breakthroughs in AI, made possible in part by increased computational power in Graphics Processing Units (GPUs), marked a new AI Spring. In the marine industry, it also marked first ever serious attempts at developing autonomous ships. Since 2018, MASS have been formally recognized by IMO, the agency of the United Nations promoting global maritime safety. Foremost on their agenda for MASS is meeting the growing need to reconcile perceived benefits of MASS implementation with growing concerns for safety (IMO, 2018).

1.1. The concept of shore-based control for unmanned ships

The future of MASS operations will aim at reducing the crew onboard and increasing the land-based coordination and control. The concept of the Shore Control Center (SCC) (also known as the Remote Control Center or Remote Operation Center) meets a growing need for centralized coordination with the tasks of monitoring, supervision, and intervention of MASS fleet operations. The scope of MASS operations for this paper includes open-ocean, short-sea, inland, urban, and mission-oriented operations, each of which set unique constraints for navigation, regulation, and safety management. Open-ocean MASS operations involve generally non-demanding navigation (e.g., Burmeister et al., 2014) whereas short-sea shipping routes involve trafficked shipping lanes containing aids to navigation and established Vessel Traffic Services (VTS) coordination (e.g., Lunde-Hanssen et al., 2020; NTNU, 2021). Inland applications apply to inland cargo vessels in inland waterways (e.g., Peeters et al., 2020a); urban applications apply to small vessels that transport passengers or goods in urban canals or waterways (e.g., MiT, 2020; Reddy et al., 2019; Valdez Banda et al., 2019; Wang et al., 2019), and data-collection Autonomous Surface Vessels (ASVs) are used in scientific research and exploratory field work (e.g., Dallolio et al., 2019; Dunbabin et al., 2009; Kimball et al., 2014). While we can conceptualize an SCC by its execution of land-based operational tasks like supervision, monitoring, and control intervention, the work undertaken by its operators will vary considerably across the range of MASS applications. For example, short-sea shipping will set more demands on navigation than open-ocean applications; urban passenger transport will set more demands on safety management that mission-oriented ASVs whose payload is just measurement equipment and data.

1.2. Definition of AI and applications to marine operations

AI has no formal definition. The first informal definition can be traced to a workshop on AI held in 1955 at Dartmouth College, commonly held to be the origin of the field. The concept was based on the "conjecture that every aspect of learning or any other feature of intelligence can be in principle be so precisely described that a machine can be made to simulate it (McCarthy et al., 2006)." This description largely holds true today, and many of the original topics under investigation at the first workshop on AI, like natural-language processing, neural networks, machine learning, and reasoning, still serve as guiding beacons for contemporary AI researchers and developers. The absence of a formal definition is not necessarily a hindrance. In a recent publication about the research state-of-the-art in AI, a committee of leading researchers stated that "the lack of a precise, universally accepted definition of AI probably has helped the field to grow, blossom, and advance at an ever-accelerating pace (Stone et al., 2016, pp. 12)." In this paper, we will examine AI applications for collision avoidance and computer vision used in the context of marine navigation. Examination of technology aspects are topical; the focus is on their implications for hazard identification, risk assessment, and human–computer interaction, with special attention afforded to the relevant theories and methods used across disciplines.

1.3. AI brittleness and the need for "humans-in-the-loop"

While there is significant potential for AI to improve safety by reducing "human errors" in marine operations, there are also many aspects of the technology that suggest new hazards. For example, computer vision based on deep learning is prone to so-called "adversarial attacks," whereby strategically modified images, often imperceptible to the human eye, deceive the algorithms and produce egregious errors (Akhtar and Mian, 2018). Real-world examples of adversarial attacks have been explored for autonomous cars (Eykholt et al., 2018) but not yet for computer vision in ships. A related problem is that of so-called "tail effects," referring to the occurrence of low-probability events that are impractical or even impossible to train as inputs into a deep learning training dataset. This introduces a subversive "tail risk;" namely, the risk arising when specific tasks like navigation are performed in an unfamiliar environment or under new conditions.

The brittleness of AI systems underscores the necessity of "humans-in-the-loop." Human operators, especially when experienced, have the remarkable ability to integrate information in unfamiliar contexts and with apparently little information (Klein, 2017). This can be thought of as the complement of AI systems, whose decision making accuracy hinges on how similar the data is to its training set. In the much-anticipated "Outcome of the regulatory scoping exercise for the use of Maritime Autonomous Surface Ships (MASS)," the IMO listed, as one of its "high-priority issues," the "Remote control station/centre." This format of operating a MASS, was, they wrote, "a new concept to be implemented… and a common theme identified in several [IMO

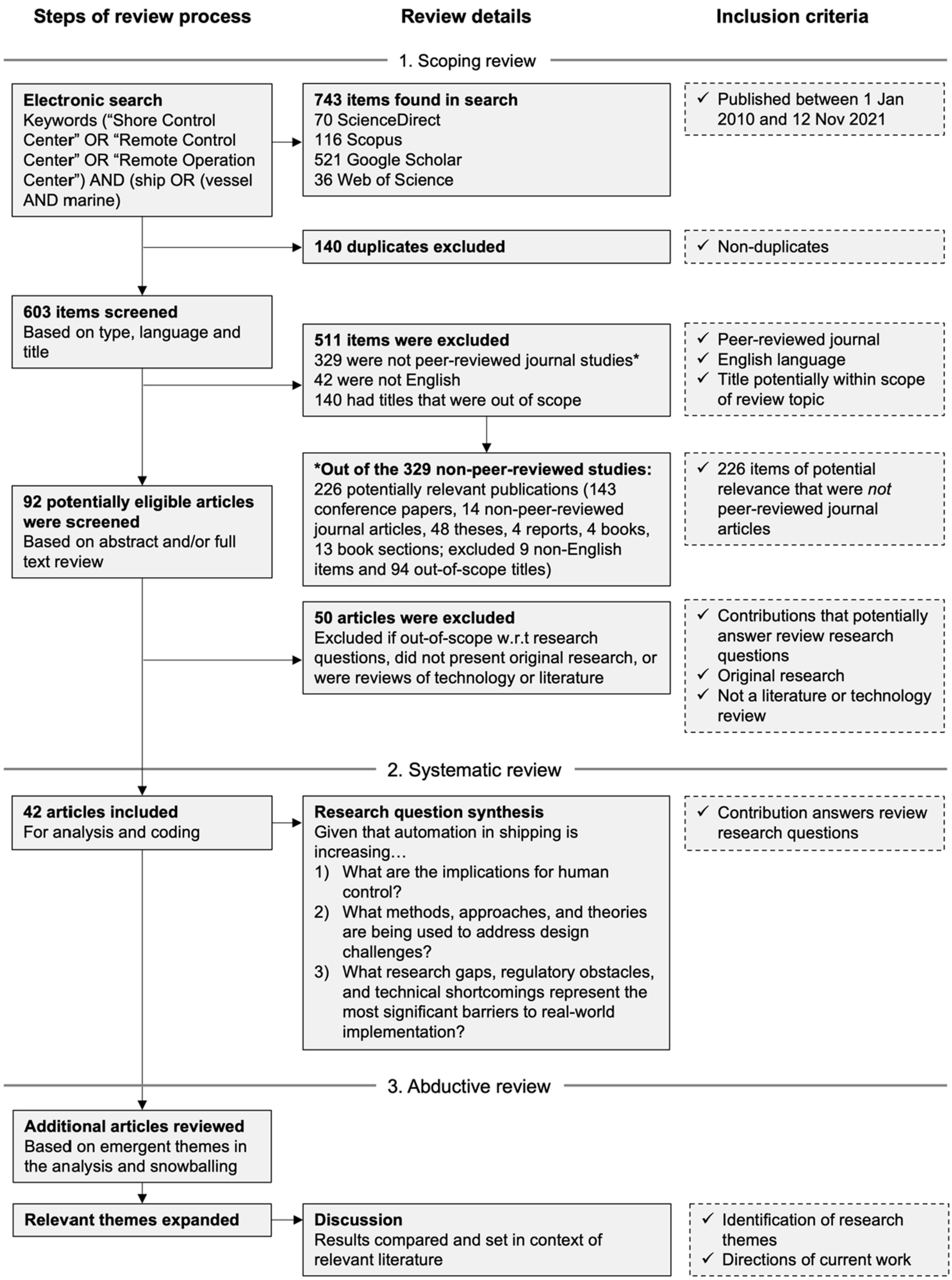

Fig. 1. Schematic showing article selection method.

Table 1
Results of systematic review.

Analysis criteria	Articles analyzed(ID can be matched to corresponding article using list in Appendix A)																				
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**
Research Question 1: How is human control being adopted?																					
Level of Autonomy taxonomies		●		●	●		●	●	●	●	●	●	●			●			●		●
Humans' roles in automated systems	●			●		●	●	●	●					●	●	●	●				●
Resilience Engineering			●														●			●	
Benefits of collaborative systems																				●	
Research Question 2: How to address design and safety?																					
Practical design approaches											●			●	●		●			●	●
Risk assessment methods																				●	
Human-computer interaction	●					●	●		●		●	●			●		●				
Empirical testing and prototypes	●			●	●		●	●	●	●	●			●			●		●		
Research Question 3: What are the outstanding knowledge gaps?																					
Risks in human-AI interaction		●		●	●	●			●		●	●		●							
Regulatory concerns		●		●	●			●			●		●		●	●					
Training for SCC operators		●	●	●	●		●	●		●			●								
Infrastructure for mixed traffic										●	●		●	●							

regulatory] instruments as a potential gap" (IMO, 2021, p. 8). Furthermore, it was outlined that the qualification, responsibility, and role of SCC operators were among the "most complex issues to be addressed." This points to the current need to articulate the concept of the SCC (called Remote control station/centre by the IMO). before researchers from diverse fields can collectively address the underlying research and design gaps.

1.4. Excessive automation and emerging risks

Recent investigations into high-profile accidents in transportation have identified interaction with highly automated systems as a root cause. Following two Boeing 737 MAX crashes in the years 2018–2019, it was revealed that the Maneuvering Characteristics Augmentation System (MCAS), designed to automatically control pitch based on sensor readings of air flow, malfunctioned without the pilots being aware that the controller existed, leading to uncontrolled dipping of the nose to counteract a non-existent stall (Nicas et al., 2019). In another accident investigation, the Tesla "Autopilot" function came under scrutiny by the National Transportation Safety Board, which surmised that "system limitations" in combination with "ineffective monitoring of driver engagement, which facilitated the driver's complacency and inattentiveness" were the probable causes of a fatal accident in California (National Transportation Safety Board, 2020, pp. 58). The task of monitoring automated systems, whether in the role of a trained pilot or passenger, is well known to be ill-suited to humans, with associated risks compounded by the need to take over control quickly (Parasuraman and Manzey, 2010; Parasuraman and Riley, 1997). Both the Boeing and Tesla examples demonstrate how poor human-AI system interaction design can lead to unacceptably dire consequences by not bringing human-in-the-loop and resilience to the fore in system design. For applications in MASS operations, which also involve control intervention in safety–critical situations, similar underlying human-system integration issues are present. Some important distinctions do exist; for example, ships have considerably slower speed, reducing the time-criticality of takeovers and simplifying interface design (Vagia and Rødseth, 2019). However, ships are also considered "under-actuated," having fewer inputs than degrees of freedom (Reyhanoglu, 1997), and may have very high inertia, requiring that navigation decisions are taken well in advance of the vessel's actual response. Controlling safety in MASS operations within acceptable risk levels will involve addressing human-AI interaction as a potential source of error, especially for time- and safety–critical interventions.

1.5. Research questions and aims

Research and development in MASS operations has increased dramatically in the last decade. In the risk sciences, significant attention has been afforded to assessing different hazard identification and risk assessment techniques for MASS operations. Designers have promoted resilience in human–computer interaction; engineers have developed and tested novel technology, and much has been written about human factors in unmanned ships. Given the emerging need for a robust human-AI interface at the SCC, as well as the contributions from multiple fields of research and the long list of obstacles slowing progression towards implementation, we organized our systematic review around three

Articles analyzed(ID can be matched to corresponding article using list in Appendix A)																					Sum
22	**23**	**24**	**25**	**26**	**27**	**28**	**29**	**30**	**31**	**32**	**33**	**34**	**35**	**36**	**37**	**38**	**39**	**40**	**41**	**42**	
●	●	●		●	●					●	●	●	●	●	●	●	●	●	●	●	*29*
●		●	●			●	●	●				●			●		●	●	●	●	*23*
		●	●	●	●	●			●			●		●	●		●				*13*
	●				●		●								●						*5*
●	●	●	●	●	●		●		●			●	●		●	●	●	●			*20*
●	●	●	●	●	●				●			●	●			●	●	●	●	●	*15*
●		●			●							●	●		●		●			●	*16*
	●								●				●					●			*15*
		●	●	●				●	●			●	●	●	●	●	●	●	●	●	*22*
	●				●					●											*11*
				●							●		●	●							*12*
●														●		●					*7*

distinct research questions. Namely, given that automation is increasing in shipping,

1. *how is human control currently being adopted in autonomous ship operations?*
2. *what methods, approaches, and theories are being used to address safety concerns and novel design challenges?*
3. *what research gaps, regulatory obstacles, and technical shortcomings represent the most significant barriers to real-world implementation?*

The review is organized as follows. Section 2 describes the review method. Section 3 describes the outcomes of the review as guided by the three research questions. Section 4 discusses the outcomes by comparing them to relevant theory and methods outside the review. Section 5 presents conclusions of the review and traces the direction of future work.

2. Method

Understanding the state-of-art within a research topic is a complex task. There are many methods available, ranging from traditional styles that adopt an inductive analysis approach and "snowball sampling" of relevant resources, to *meta*-analysis approaches that aim to synthesize with the help of statistical tools an entire field of research, often citing hundreds of studies. In this paper, we adopted a systematic, comprehensive review approach that can be considered somewhere between traditional and *meta*-analysis review approaches. The systematic review has three virtues: it is reproduceable, introduces little bias, and it is transparent. Systematic reviews come in no universal form, although one commonly used guideline called "Preferred reporting items for systematic reviews and *meta*-analyses" (PRISMA) calls for a degree of standardization among researchers (Moher et al., 2009). We adopted PRISMA guidelines in our review.

The scope of the review was guided by the need for state-of-the-art research, not simply a state-of-the-art in implementation. We searched specifically for research about remote and shore-based control of autonomous ships because it conceptually represented the integrated MASS system, including not just elements of automated navigation but also of human control and emergent safety properties. We chose our search keywords ("Shore Control Center" and various alternatives) because these keywords represented this integrated system emerging in the wake of MASS development. The concept is also used by different fields in different ways, agnostic to any one research discipline. Moreover, the concept is relevant to all the research questions we set out to investigate. The method is divided into three phases: a scoping review, a systematic review, and an abductive review (Fig. 1).

(i) **Scoping review**: electronic search on the research topic followed by a series of screenings. The screenings occurred in two stages:
   1. Initial screening: based on title, type, and language: only titles that reflected the scope of our research questions were included, as were peer-reviewed journals, and English-language studies.
   2. Abstract screening: based on consensus of the two authors who independently analyzed the abstracts of studies passing initial screening.

(ii) **Systematic review**: full-text analysis by first author and coding in text analysis software NVivo (NVivo, 2020) according to the three research questions and relevant bibliographic details.
(iii) **Abductive review**: comparison to literature outside of the analyzed studies and expansion of the topics analyzed in the results.

2.1. Scoping review

Four databases were searched: ScienceDirect, Scopus, Web of Science, and Google Scholar. This ensured a broad indexing of academic databases, although it also resulted in many duplicates. The date range was set to 1 Jan 2010 to 12 Nov 2021. The search was restricted to the keywords "Shore Control Centre," "Remote Operation Centre," and "Remote Control Centre," plus an additional phrase to narrow the scope to marine applications ("AND Ship OR Vessel AND Marine").

Note that using double quotations in search phrases (e.g., "Shore Control Centre") in ScienceDirect and Scopus specified a search with those words together, but also included variants on punctuation and regional spelling and plural forms as well as variations on capitalization. For example, "Shore-Control Centers" and "Shore Control Centre" appeared together in the same search. For Web of Science and Google Scholar, the regional spelling variants had to be specified using the search "Shore Control Centre" AND "Shore Control Center" (but maintained all same flexibility in plurals, hyphenation, and capitals as in the other search database rules).

The initial screening was straightforward: duplicates and non-English language contributions were removed as were any publications from non-peer-reviewed journals. Note that excluding all technical reports and conferences papers potentially excluded valuable contributions, presenting a limitation in our review method. On the other hand, this exclusion criterion ensured a consistent standard of research contributions in the analysis. Titles of all remaining peer-reviewed journal articles were screened for those that were obviously out-of-scope.

The second-stage screening of articles required analyzing their abstracts and reaching a consensus on whether they should proceed to full-text review. Analysis was conducted by the two authors. Eligibility was based on whether the study potentially contributed to either one of the

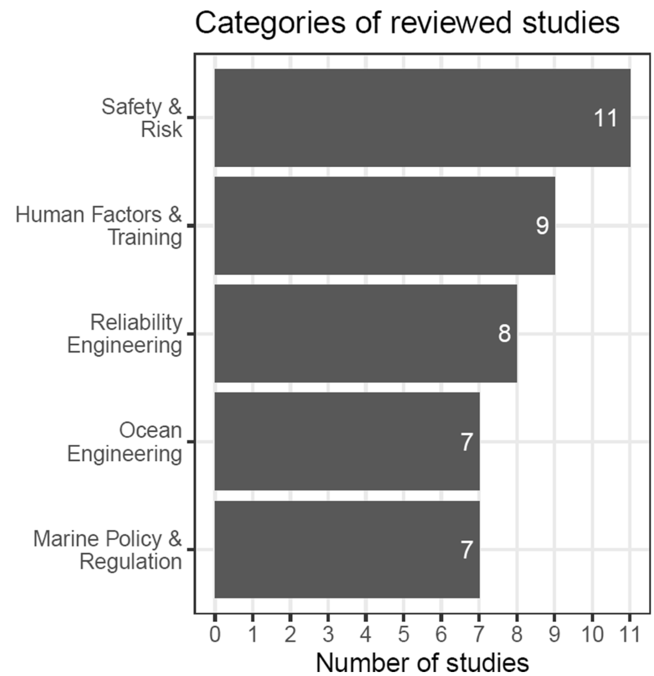

Fig. 3. Categories of all analyzed studies.

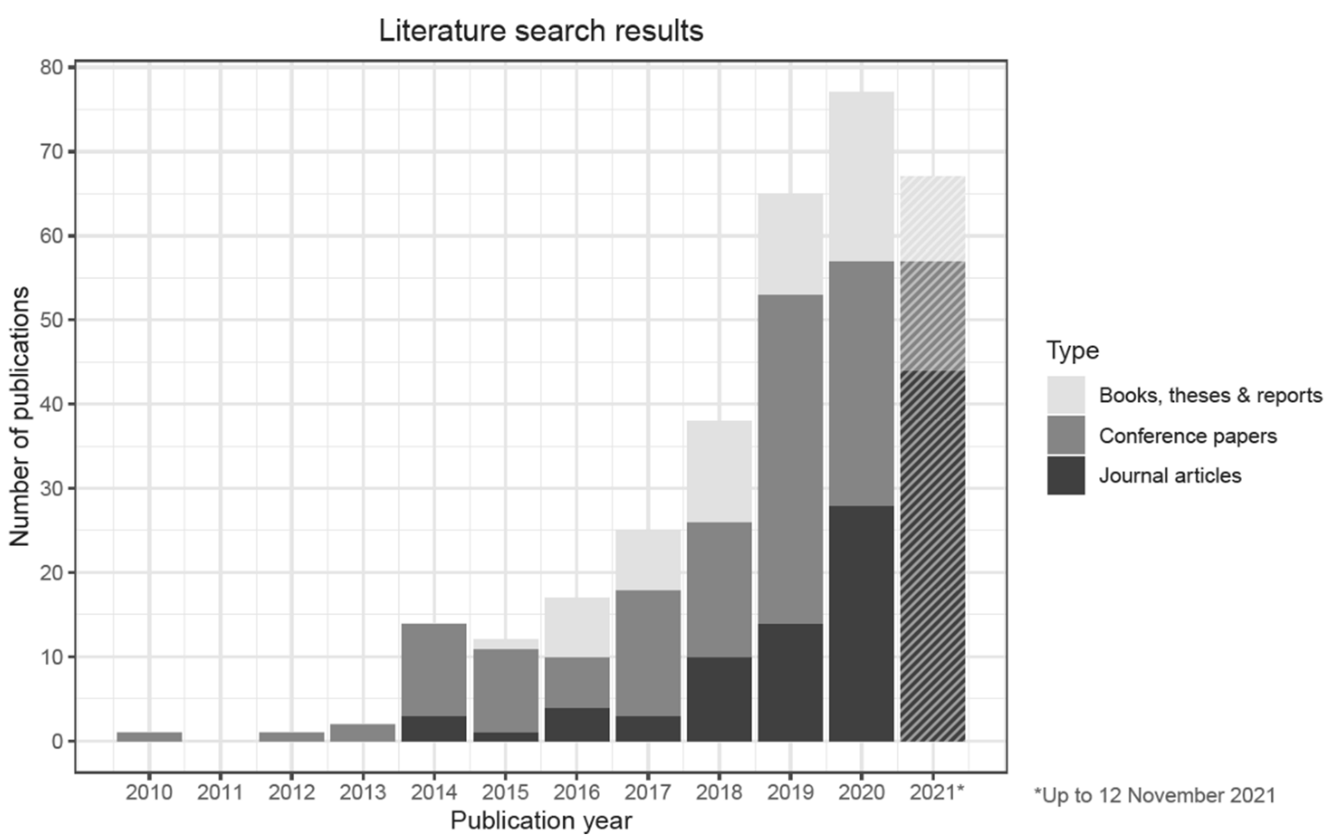

Fig. 2. Number of SCC studies published by year.

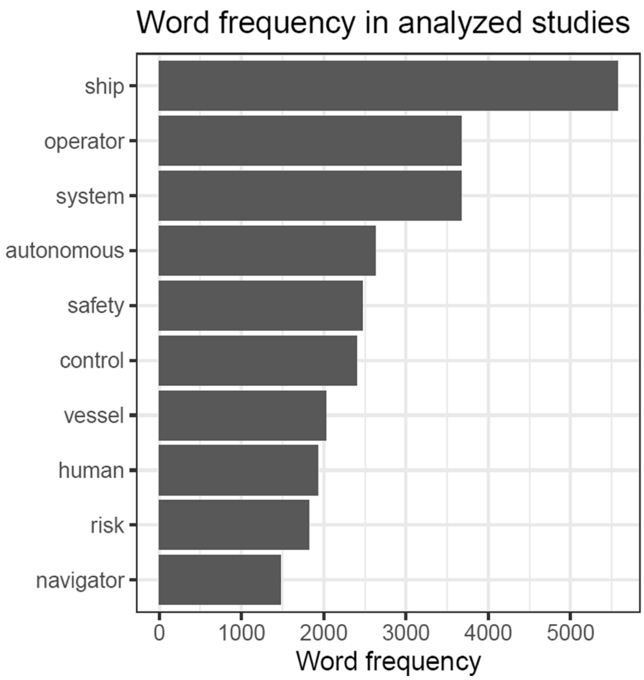

Fig. 4. Salient themes by word frequency in all analyzed studies.

three research questions defined in the Introduction. This was not always obvious from just the abstract; Grote et al., 2014 when in doubt, we included the study on the basis that the subsequent full-text analysis would confirm whether the study should be included. Also, only studies presenting original research were included, excluding opinion papers and review studies.

During the second-stage screening, we also made some additional scoping constraints to help converge our study subset. We briefly note these constraints here. We defined the scope to studies including any integrated system consisting of one or more MASSs and an SCC. From here on in we refer to such systems as just "MASS systems." The context was broad, including information needs, design approaches, systems design, complex socio-technical perspective, and elements of human–machine interaction related to remote control and intervention (from safety, risk, regulatory, technological, and maritime training perspectives) from shore-based infrastructure or from other surface vehicles. Importantly, this excluded work about UAVs and ROVs and papers focused on business cases or economics, cyber-security, law, cybernetics, insurance, or discussions about taxonomy for autonomous vessels. Additionally, we excluded works focused on just one component (e.g., communication or motion control) that only peripherally mentioned the human control and human interaction elements. The guiding principle was that the studies as a whole should reflect state-of-the-art research on MASS systems and not a state-of-the-art on its implementation.

2.2. *Systematic review*

Following the PRISMA guidelines, our main aim was to qualitatively synthesize the literature in the context of the defined specific research questions in a way that was reproduceable, unbiased, and transparent.

2.2.1. *Review protocol, data extraction & coding*

A review protocol specified how the studies were screened during full-text review against eligibility criteria and how information was extracted for answering the research questions.

Studies were screened for information that contributed to any one of the three defined research questions. This was done by coding the studies in the text analysis software NVivo (NVivo, 2020). Coding for the

Table 2
LoA taxonomies adopted in the reviewed studies.

LoA taxonomy	Description of maximum LoA for current operational concepts	Article ID
IMO (2018)	**Remotely controlled ship without seafarers on board (Degree 3 out of 4):** The ship is controlled and operated from another location. There are no seafarers on board.	[2, 4, 5, 8, 9, 10, 11, 13, 22, 23, 33, 35, 38, 39, 42]
NFAS (Rødseth, 2017)	**Constrained autonomous (Level 7 out of 8):** This is assumed to be a common mode for unmanned ships operating in relatively benign environments. The ship can operate fully automatic/autonomous in low traffic and non-restricted waters, until it sees problems that it cannot solve itself. Then it can call on shore control to resolve any problems. This reduces ship system complexity, while optimizing the benefit of having backup functions on shore.	[12, 19, 24, 37]
Utne et al. (2017)	**Semi-autonomous operation or management by exception (Level 3 out of 4):** The system automatically executes mission-related functions when response times are too short for human intervention. The human may override or change parameters and cancel or redirect actions within defined timelines. The operator's attention is only brought to exceptions for certain decisions (human-supervisory control).	[27, 34]
Lloyd's Register, (2016)	**"Active" human in the loop (Autonomy Level 3 out of 6):** Decisions and actions at the ship level are performed autonomously with human supervision. High impact decisions are implemented in a way to give human operators the opportunity to intercede and override them. Data may be provided by systems on or off the ship.	[26, 41]
Central Commission for Navigation on the Rhine (CCNR, 2021)	**Conditional to High Automation (Level 3–4 out of 5):** The sustained context-specific performance by a navigation automation system of all dynamic navigation tasks, including collision avoidance, with the expectation that the human helmsman will be receptive to requests to intervene and to system failures and will respond appropriately (no human intervention needed for High Automation). Applies to inland MASS only.	[16, 32, 40]
Combination / dynamic LoA	Shifts between generalized LoA levels, especially under **"adaptive autonomy"** as defined by Sheridan (2011) whereby "allocation of control function (to human or computer) changes with time to accommodate changes in the conditions of either the physical environment or the human" (p. 662).	[7, 21, 36]

three research questions also helped to identify likenesses among studies across more granular dimensions – dimensions that, when assigned a title, became the sub-categories that structured the presentation of the Results. There are, namely, four such sub-categories for each research question presented in the Results section.

The following information was coded during analyses of the studies:

(i) Research Question 1: How is human control currently being adopted?
  a. Level of Automation (LoA) taxonomies
  b. Humans' role in MASS operation
  c. Resilience Engineering
  d. Benefits of collaborative systems
(ii) Research Question 2: What methods, approaches, and theories are being used to address design challenges and improve safety?
  a. Practical design approaches
  b. Risk assessment methods
  c. Human-computer interaction paradigms
  d. Empirical testing and prototypes
(iii) Research Question 3: What research gaps, regulatory obstacles, and technical shortcomings represent the most significant barriers to implementation?
  a. Risk assessment methods
  b. Regulations
  c. Design for human-AI interaction
  d. Infrastructure to support mixed traffic

The coded data relevant to the research questions are summarized in Table 1.

Studies were also screened for bibliographic and descriptive details. The purpose of the former was to tabulate a list of all analyzed studies, which is included Appendix A. The purpose of the latter was to broadly characterize the contributing publications in terms of journal scope (which we sought to bin into categories) and in terms of salient themes (which could be represented by textual analysis of the most frequently used keywords). The following bibliographic and descriptive information was coded during analyses of the studies:

Table 3
Humans' roles in MASS system.

Category	Description	Article ID
Active	**Continuous monitoring and decision making support:** operator's role is active and engaged in all operational phases; the operator is in control of the ship, either through direct or indirect remote control. On-board roles were often specified, including emergency handling, active maintenance, lookout / watchkeeping, cargo loading and unloading.	[1, 7, 9, 15, 16, 17, 24, 29, 30, 34, 40, 42]
Backup	**Monitoring and control intervention:** operator's role is characterized as "backup" to the AI system; control interventions, or takeovers, can occur when the operator takes over control from the AI system, either on their own initiative or from the prompting of the AI system itself. Monitoring is mostly continuous; the operator is never far from the control position if left unattended; emphasis placed on timely emergency or contingency response.	[4, 6, 8, 21, 25, 37, 41]
Passive	**Supervision and assistance:** operator's role is characterized mainly by passive supervision; the operator can leave the control position and is alerted by the AI system if they are needed. Instead of handling situations as they arise, emphasis is on planning how to resolve situations before intervention is needed.	[14, 22, 28, 39]

(i) Bibliographic information (author, year, title, and journal name)
(ii) Descriptive information:
  a. Journal scope (Safety & Risk, Human Factors & Training, Reliability Engineering, Ocean Engineering, Marine Policy & Regulation)
  b. Salient themes (computed from keyword frequency).

2.3. Abductive review

The abductive review is presented in the Discussion. Here "abduction" refers to the process of forming an explanation for a phenomenon. It presupposes that the explanation we present are grounded at least partly in the works we analyzed (and partly in works outside the review) and do not simply represent our conjectures or guesswork. The term was coined by the philosopher Charles Peirce and has inspired interpretations for its use in modern science (for example, Paavola, 2006; Swedberg, 2014). We adopt the interpretation that forming an explanation is a principled process. We present our Discussion accordingly, in five sub-sections that expand upon central themes uncovered in the analysis. Inferences are drawn transparently to the Results as well as to literature outside the review's scope. This includes a diverse range of topics, including human-AI interaction for autonomous cars, ethnographic studies of control room work, the growing use of simulation in human-AI interaction studies, as well as the role of safety management in striking a balance between human and machine autonomy. In the abductive review we also reviewed studies cited in the system review in a so-called "snowballing" approach to scoping relevant literature.

3. Results

3.1. Electronic search

The electronic search resulted in 743 studies. This included, in respective order, 70, 116, 521, and 36 for ScienceDirect, Scopus, Google Scholar, and Web of Science. This included research and review articles, books, conference papers, and reports published from 1 Jan 2010 to 12 Nov 2021. After removing duplicate studies, 603 remained. Each of these was screened by title and by language, removing those studies that were clearly out-of-scope and those in a language other than English. The remainder was 318 studies, comprised of 92 peer-reviewed journal articles and 226 other publications (conference papers, books, theses, and reports). Fig. 2 presents a plot of these 318 studies organized by date and frequency of occurrence. It shows an abrupt start in the years 2013–2014, followed by steady publication rate through to 2017, after which publications increase markedly in 2018 through to 2020. The year 2021 saw a relative increase in the number of journal articles and decrease in amount of conference papers – likely a result of the COVID-19 pandemic. Note that the first publication describing the concept of a MASS system was in a 2010 paper by Im and Seo (2010), who described a "free running" or "automatically navigating" ship controlled from an SCC. The concept of "autonomous ships" with integrated control from a "shore control center" was first formally described by Rødseth et al. (2013).

Excluding all conference papers, non-peer-reviewed journal articles, theses, reports, books, and book sections, we were left with 92 potentially eligible peer-reviewed journal articles. After scanning the abstracts of these 92 articles, a further 50 were excluded. Note that it was not always obvious from the abstract whether a study should be included; in these cases, the full text was scanned, often after discussion between the co-authors until consensus was reached. Full-text analysis and coding was done for a total of 42 studies (Appendix A).

Categorizing the 92 articles by journal scope, we found that "Safety & Risk" and the related category of "Reliability Engineering" together accounted for 45% (n = 19) of the total studies analyzed (Fig. 3). A text frequency query, computed in NVivo (NVivo, 2020), showed that the word "safety" appeared among the top five most frequent words (excluding words with fewer than three letters and including stemmed words) (Fig. 4).

3.2. Research question 1: Given that automation is increasing in shipping, what are the implications for human control?

In this section, we explore the question of why human control is needed despite increasing automation. Drawing exclusively from the reviewed articles, we examine four relevant themes: (1) Level of Automation (LoA) taxonomies, (2) humans' roles in MASS systems, (3) Resilience Engineering (RE), and (4) the benefits of collaborative human-AI systems. Together, these four themes offer a cross-section of the reviewed studies that explain what at first glance seems to be a paradox: that designing for increased automation in shipping hinges upon design for human control.

3.2.1. Level of automation taxonomies

Most of the articles analyzed (twenty-nine out of forty-two), adopted a Level of Automation (LoA) taxonomy. LoAs are a way of describing how autonomous a system is by way of a linear integer scale, usually from 0 indicating full human control to some maximum number representing full machine autonomy. All articles adopted the position that full autonomy was not feasible in the scope of current research. That is, even though the terms "autonomous ship" was used, it was understood to be a "highly automated ship" involving some level of mixed human-AI control as defined by an LoA taxonomy. To illustrate this, Ramos et al. (2019, 2020b) express that full autonomy is "not expected in the near future" and Huang et al. (2020) take the stance that it is "still unrealistic." In Hannaford and Hassel (2021), the authors surveyed forty-two licensed seafarers on the matter, of whom two-thirds responded "it will not be safe to operate vessels in the future as fully autonomous" (p. 13). Regardless of what LoA taxonomy was being used, all studies positioned themselves one or two steps below the maximum number to show that the ship was highly autonomous, albeit not autonomous enough to exclude humans. Specifically, MASSs still relied on humans enough to merit the SCC in the system design.

By far the most commonly referenced LoA taxonomy in the papers analyzed was that proposed by IMO (2018). Fifteen out of the twenty-nine papers that positioned themselves within an LoA framework used this definition (Table 2). The second most commonly cited was NFAS (Rødseth, 2017), with four articles. Other LoAs used included one by Utne et al. (2017), Lloyd's Register of Shipping (Lloyd's Register, 2016), and the Central Commission for Navigation on the Rhine (CCNR, 2021) (the latter applying to inland MASS only). Three articles, while acknowledging the existence of various LoA taxonomies on offer, opted not to adhere to just one of strict definition. Instead, these studies generalized the MASS system as a combination of machine and human control, dynamically shifting with time. For example, Thieme et al., (2018) referenced several LoAs but settled on the general term "adaptive autonomy" put forward by Sheridan (2011) to avoid adherence to one particular LoA taxonomy. For a detailed review of LoA taxonomies, the reader is referred to Vagia et al. (2016).

3.2.2. Humans' roles in MASS systems

Just over half the analyzed papers (twenty-three out of forty-two) described specific roles human operators play in the MASS system. The studies were unanimous in their assertion that operators would be responsible for ensuring safety of the ship and any passengers onboard, and for ensuring protection of the surrounding environment. The ability for humans to oversee several ships at once was also described as a common feature of SCC work, as was the ability to plan voyages and take over control of the automated system.

Still, there were some notable differences in how humans' roles were described. In our review, we identified three broad categories of operator roles: (1) Active, (2) Backup, and (3) Passive. In Table 3 we characterize these three roles and list their corresponding studies. In approximately half the studies (twelve out of twenty-three), humans were identified as having an "active" role, characterized by "continuous monitoring and decision making support." The remaining role descriptions were underpinned by "backup" (seven studies) and "passive" roles (four studies). One common feature to all three roles, as noted in the studies, was their susceptibility to change, dependent as they were on technology development and operational elements (busier ports, for instance, require more active operator engagement). Still, each study could be mapped to just one of the three categories. This categorization was usually based on the case studies presented or on underlying assumptions about MASS system design. For example, Rødseth et al. (2021), while acknowledging the wide range of possibilities for human participation, narrow down the operators' primary role in the MASS system as one of backup to the automation. "It is expected that most autonomous ship systems will operate with continuous supervision from an RCC [Remote Control Center]," they write, specifying that operators need to "intervene when the automation is incapable of maintaining control" (p. 5). Wróbel et al. (2021), on the other hand, emphasize that a more active role related to diagnosing problems is needed to meet the expectation that problems can be "rectified" in a timely way. Meanwhile, others highlighted the need for planning and logistics, especially as more ships fall under command of a single SCC (e.g., Reddy et al., 2019).

Note that the categories below are independent of the studies' chosen LoA taxonomy. To illustrate this, some studies described operators' tasks as becoming more passive with higher automation (e.g., Zhou et al., 2021), while others argued that higher automation comes with the need for more active coordination among different roles (e.g., Huang et al., 2020; Relling et al., 2021).

Some studies stood out in their treatment of specific roles in the MASS system. Eriksen et al. (2021), for example, highlight that maintenance activities are critical to a ship's operability and rely on manual repair. Kooij and Hekkenberg (2021) also allude to the importance of repair activities onboard ships, pointing out that, "in contrast to [highly automated] cars and aircraft, the operators on ships do much more than

Table 4
Design approaches adopted in the reviewed studies.

Goal-based design approach	Description	Article ID
Human-computer interface design	Functional requirements are defined in terms of tasks, task goals, and operational objectives. Methods for decomposing operational objectives into distinct task goals included Task Analysis (TA) and Information-Decision-Action (IDA) (described in more detail in Section 3.3.3).	[11, 15, 17, 22, 24, 27, 29, 31, 34, 35, 37]
Safety design	Functional requirements are defined in terms of acceptable risk levels. Although this can take many forms, the over-arching purpose in generally the same: to identify hazards, estimate their occurrence, predict associated risks, and design safety controls for mitigating these risks (see Section 3.3.2 for a full list of risk analysis methods). Often referred to as "risk-based design."	[20, 22, 23, 24, 25, 26, 27, 29, 31, 34, 35, 37, 38, 39, 40]
System integration	Functional requirements stem from integrating components of a complex system. This may include diverse elements like human–computer interaction, operator tasks, and risk-based design, and emphasizes consistent modeling techniques across the board. Often referred to as "systems thinking," or "systems engineering" approaches to design.	[14, 21, 31]

'just' navigating the vehicle" (p. 2). In this sense, some human roles can be seen as holdouts to increasing automation, which include not just machinery maintenance and repair, but also cargo loading/unloading and mooring at port.

Saha (2021) draws similarities between the role of MASS operators and those of Officers of the Watch (OOW) on conventional ships, in that they are both "responsible for monitoring the ship and intervening if needed" (p. 10). Hannaford and Hassel (2021), in contrast, hesitate to describe human roles at all, on the grounds that "it is too early to identify the amount of crew and the specific roles that will be needed because the crew has not been replaced yet, and not enough research has been conducted or published" (p. 15). Yoshida et al. (2021) argue that overall responsibilities of SCC operators will increase when compared to navigators aboard conventional ships.

3.2.3. *Resilience Engineering*

Resilience Engineering (RE) is a paradigm that aims to balance productivity with safety in complex systems using safety control techniques where risks are managed over time. Instead of focusing on human errors and inferring safety from accident investigations, RE positions humans as the source of resilience in a complex system and places the emphasis on safety controls. This way of thinking sheds light on the implications for human control in the MASS system, even before we have such systems implemented. Almost one-third of the analyzed papers (thirteen out of forty-two) adopted a design framework that fit within the RE paradigm.

Here we present the relevant studies in terms of four broad principles at the core of RE. These are (1) focus on human as a source of resilience rather than as a source of error, (2) emphasis on safety controls and dynamic risk indices, (3) use of systems engineering-based risk assessment technique like System-Theoretic Process Analysis (STPA), and (4) approaches to safety science that consider human and organizational interaction elements in addition to more straightforward component failures.

Starting with the first principle, Ahvenjärvi (2016) directs attention towards the roles of software developers of the MASS system on the principle that here, too, can resilience be built in given that we design for front-line operators' flexibility and problem solving abilities. In line with the second principle, Thieme and Utne (2017) apply the "resilience-based early warning approach to development of indicators" (REWI) method to the MASS system, focusing on handling accidents, incidents, and unexpected events from an organizational perspective. Utne et al. (2020) elaborate on the "supervisory risk control" concept, showing that the system's resilience goes beyond conventional safety controls by virtue of its ability to learn, adapt and improve over time. Reddy et al. (2019) highlight the need for a "systematic, traceable, and holistic assessment" of safety controls which shares some of the organizational and managerial elements discussed in Thieme and Utne (2017) and Utne et al. (2020). Peeters et al. (2020b) designed their prototype specifically to be able to measure performance indices over time; similarly, Wu et al. (2021) designed a real-time safety index algorithm based on historical sailing voyages along the same route.

For the third RE principle, Wróbel et al. (2018) were the first to apply STPA to MASS system risk assessment. In the same year, assessing the applicability of sixty-four risk assessment methods to MASS systems, Thieme et al. (2018) promote STPA as among the most suitable candidates, in part for its ability to identify human–machine interaction risks and propose risk reduction methods. Valdez Banda et al. (2019) also used STPA, in their case to inform early-stage risk-based design of an autonomous urban passenger ferry. Zhou et al. (2021) also apply the STPA approach to MASS system design, with a special focus on identifying both safety and cyber-security hazards and their inter-related safety controls. In line with the fourth RE principle, Ramos et al. (2020a, 2020b) developed the "H-SIA method" to enable analysis of the complex system as whole, rather than modeling how component failures can propagate. Ramos et al. (2019) take a similar approach for human task analysis in MASS operations, linking the success of MASS operators' tasks to latent socio-technical factors like training and organizational culture. Overall, the studies that adopted RE principles placed humans' roles in the MASS system as central to ensuring system safety.

Table 5
Risk assessment methods adopted in the reviewed studies.

Name	Description	Article ID
BNs	**Bayesian Networks**: used to update prior beliefs about system behavior to model risk over time in an interactive network.	[27, 35, 38, 41]
FMEA	**Failure Modes and Effects Analysis:** a classic risk assessment method combining severity, likelihood of failure mode, and detection rate to compute component or system level failure probabilities.	[38, 42]
H-SIA	**Human-System Interaction in Autonomy:** "provides a framework for analyzing autonomous ship operation as an entirety, rather than each agent separately" (Ramos et al., 2020b, p. 1).	[24, 34]
REWI	**Resilience-Based Early Warning Approach to Development of Indicators**: aims at determining organizational cap- abilities to handle unexpected situations and producing quantifiable safety indicators.	[25]
STPA	**System-Theoretic Process Analysis**: used for identifying hazards and revealing causal factors especially for novel and complex systems for which there is little experience or empirical data.	[26, 27, 31, 39]
THERP	**Technique for Human Error Rate Prediction:** risk assessment focused on quantifying human errors; used in the quantitative analysis of human reliability in human factors domain.	[35]
RCM	**Reliability Centered Maintenance**: stepwise method developed in the aviation industry to optimize maintenance management and improve reliability of safety–critical systems.	[30]
FSA	**Formal Safety Assessment:** a process developed by IMO in the wake of Alpha Piper disaster; consists of systematic hazard identification, risk assessment and control, cost-benefit analysis, and safety recommendations.	[40]
CPA	**Closest Point of Approach:** risk is assessed based on proximity of other objects, or from deviation of a target from its planned path; risk is a considered as a function of the target's spatial–temporal properties.	[20, 22, 23]

3.2.4. *Benefits of collaborative human-AI systems*

Collaborative human-AI systems are a way of framing increasing automation as an effort to enhance both human and machine roles by tailoring both according to the strengths and weaknesses of the two counterparts. The purpose of increasing automation, in this sense, should not be seen simply as an effort to reduce occurrence of human errors. Rather, collaboration between human and machine should lead to greater system performance than could be achieved by either counterpart working in isolation.

Five studies out of the forty-two analyzed identified benefits of human-AI teaming that went above and beyond human error reduction. Wu et al. (2021b, 2021a), for example, develop interface designs for MASS operators that display a comparison of current navigation activities relative to those of successful historical voyages on the same route. The AI computes large amounts of data and alerts the operator only when a safety index is exceeded – when the ship deviated from its course. At this point, the operator can use the output as a heuristic for decision making (does the course deviation warrant intervention?). In this way, a synergy between the computational power of the AI system and the decision making capacity of an expert operator emerges. To consider another example, Thieme and Utne (2017) frame shore-based decision support, control, and operational logistics in terms of "dynamic safety performance monitoring." This view presumes that primary navigation tasks are computed by AI while humans attend to the higher-order tasks involved in safety management: a partnership that augments operators' role of safety–critical decision making. Similarly,

Table 6
Human-computer interaction paradigms adopted in the reviewed studies.

HCI paradigm	Description	Article ID
Cognitive modelling	**Situation Awareness**: a concept used in a general sense to describe an operator's semantic apprehension of their surroundings; often specifically refers to Endsley's cognitive model called of the same name (Endsley, 1995).	[1, 6, 11, 17, 22, 26, 35, 39, 42]
	Information, Decision, Action in Crew context: a framework introduced by Chang and Mosleh (2007) to model cognitive processes of crew in response to accidents in nuclear power plant operations.	[24, 34, 37]
	Task Analysis: "collective noun used in the field of ergonomics, which includes HCI [human–computer interaction], for all the methods of collecting, classifying, and interpreting data on the performance of systems that include at least one person as a system component" (Diaper, 2004).	[15, 24, 34, 37]
Embodied cognition	**Sense-making:** an "enactive" approach in cognitive science (distinct from "functionalist" cognitive paradigms defining the body only in terms of input and outputs); "cognition as sense-making is the exercise of skillful know-how in situated and embodied action" (Varela et al., 2016, p. xxvi). Also referred to as "ship sense" and "harmony."	[1, 6, 11, 12, 17]
	Human factors: examines latent causes of error or factors that improve system interactions in a socio-technical system. This may include cognitive functions and sense-making, but also other phenomena like vigilance, workload, trust, and fatigue.	[6, 7, 9]

Huang et al. (2020) and Ramos et al. (2019) frame AI-operator teaming in terms of enhancing practical control-related tasks like object detection and coordinating collision avoidance, rather than in terms of wholesale substitution of human control.

3.3. Given that automation is increasing in shipping, what methods, approaches, and theories are being used to address design challenges?

The forty-two studies analyzed revealed a broad range of methodological approaches. Each of these approaches represented a different investigative angle to the same underlying problem; namely, how can design challenges emerging from increased automation in shipping be addressed. Here we break down this methodological dissonance into four categories representing common strategies among the various theoretical frameworks. Each of these categories are analyzed separately below and include: (1) practical design approaches for human–computer interaction, (2) risk assessment methods used in risk-based design, (3) different human–computer interaction paradigms adopted, and (4) empirical testing or data collection efforts. These four categories emerged during the analysis, encompassing the wide array of perspectives contributing to the same topic. This "common denominator" approach to summarizing the variety of methods on offer provides a handle on the relevant theoretical foundations adopted in the contributing studies.

3.3.1. Practical design approaches for human–machine interaction

Almost half of the studies (twenty out of forty-two) adopted a practical approach for designing human–machine interaction elements of the MASS system. We summarize these approaches here, characterizing them broadly within three sub-sets of goal-based design (Table 4). Here we use the term "goal-based design" to describe the general process of mapping high-level system objectives to functional requirements. We identified three goal-based design sub-sets: (1) human–computer

Table 7
Studies that present empirical observations about MASS systems.

Source of observation	Description	Article ID
Field test	A prototype unmanned inland cargo vessel (scale model, length 4.8 m) was tested in inland waters to demonstrate its feasibility.	[19]
Field test	The SCC counterpart to [19] was tested for operator situation awareness. The researchers aimed to "stress test" the SCC prototype and investigate whether eye-tracking could provide insights into operators' SA (n = 2 participants).	[17]
Field test	Sea trials were undertaken for an "autonomous cargo test ship" in a test area outside of Wuhan, China; the ship was controlled remotely from a testing center in Netherlands.	[35]
Simulator test	First test program in a full-mission bridge simulator re-configured as an SCC, designed to investigate operators' situation awareness (n = 6 participants, all experienced seafarers).	[1]
Simulator test	Researchers designed and conducted an experiment to study VTS-MASS interactions. Testing was done in a full-mission bridge simulator (n = 24 participants; half experienced seafarers, half control group).	[10]
Simulator test	Commercial full-mission bridge simulator used to measure effect of a "guidance-support system" designed by the researchers to enhance collision avoidance performance (n = 36).	[23]
Survey	Participants were asked to evaluate the extent to which they thought that sixty-six different Knowledge, Understanding & Proficiency requirements (KUPs), set out in the STCW convention, were relevant for autonomous shipping. Responses were collected on a Likert scale (n = 109, all seafarers with a navigation license).	[4]
Survey	Participants were asked to evaluate the relevance of seventy-five human factors for SCC operators (n = 32, seafarers).	[9]
Survey and interview	Participants were asked about employment, training, safety, and feasibility of MASS systems, with results collected on a Likert scale (n = 42, all licensed navigators). Subject Matter Experts were also interviewed about similar topics (n = 7, variety of background expertise).	[5]
Workshop	Held with VTS operators in Norway; participants generated ideas about the role VTS may play in the MASS system, with special focus on regulatory challenges (n = 26).	[14]
Workshop	Series of four workshops held with experts to identify hazards and ideate safety control functions for an autonomous urban passenger ferry in early-stage design (Workshop 1–3: n = 8, Workshop 4: n = 7; participants were a variety of subject matter experts with some return participants; n = 20 total individuals).	[31]
Workshop	Series of two expert workshops to assess risks in the MASS system (Workshop 1: n = 12, variety of subject matter experts; 2: n = 8, return participants).	[40]
Interview	Semi-structured interviews with expert informants about competence requirements for SCC operators (n = 10, subject matter experts from academia, regulatory agencies, technology companies, and other relevant backgrounds).	[8]
Interviews and questionnaire	Interviewed seafarers (n = 10) and subsequently had participants watch a video of a simulator running scenarios (n = 25 trials); participants also filled out a NASA-TLX questionnaire for perceived workload.	[7]
Focus group	Interviewed experienced nautical sciences instructors (n = 3) about situation awareness in SCC operations in light of current STCW conventions on training and competency requirements for seafarers.	[11]

interface, (2) safety, and (3) system integration.

Eleven studies adopted a goal-based design approach oriented towards "human–computer interface design." Peeters et al. (2020b), for example, design and build an SCC that is integrated with a model-scale MASS prototype. Their approach starts with coming to terms with the various "activities" of operators, which take shape from assignment of operational objectives. Similarly, Yoshida et al. (2020) propose functional requirements for an interface designed to support operator situation awareness based on a "human behavioral model" of the SCC. Kooij and Hekkenberg (2021) focus on what tasks need to be replaced when moving seafarers off MASSs, leading to suggestions for a crew-reduction implementation plan. This also leads to functional requirement in terms of how many operators are onboard a given MASS and at what times. For example, the latter authors propose, "Replacing mooring, deck maintenance, bunkering, administration, and port supervision decreases the required crew to four crew members" (p. 15).

Fifteen studies adopted a goal-based design approach oriented towards "safety design." Wu et al. (2021a) propose a safety indicator system intended specifically for SCC operators and demonstrate it using data collected from the log files of sixteen voyages completed by a research vessel. The technique uses a supervised machine learning algorithm in a "black box" approach to obtain what is essentially a real-time safety indicator calculator active during a ship's voyage. Wu et al. (2021b) do something similar, based on historical data collected from VTS logs. Several other studies systematically identified hazards for the purpose of developing functional requirements for safety controls (Bolbot et al., 2021; Chang et al., 2020; Thieme and Utne, 2017; Wróbel et al., 2018; Zhou et al., 2021).

Three studies adopted a "system integration" approach to goal-based design. Relling et al. (2021) propose ways to integrate VTS with MASS traffic coordination, calling their approach a mixture of "systems thinking and participatory design thinking." Rødseth et al. (2021) argue for a systems approach to MASS design that develops functional design requirements from a range of related elements under a so-called "Operational Design Domain." This includes elements like operational complexity, human–computer interaction, and geographic location. The design work presented by Valdez Banda et al. (2019), in considering functional requirements over the entire life cycle of a MASS, is also exemplary of a systems engineering approach.

The three goal-based design approaches identified were not mutually exclusive and were often combined. Ramos et al. (2019, 2020a, 2020b), go to great lengths to describe operator tasks and link them to risks associated with MASS operational design. Utne et al. (2020), Huang et al. (2020), and Zhang et al. (2020) also combine elements of interface and safety design, in their cases mapping cognitive elements related to task handling (like perception, action, and prediction) to hazard identification and failure probabilities and their associated risks. Valdez Banda et al. (2019) combine elements of all three approaches, adopting risk-based design within the classic naval architecture "design spiral" (Evans, 1959), and adhere to systems engineering principles for life-cycle design. Yoo and Lee (2021) develop requirements for VTS and SCC operator interface designs based on collision avoidance risk assessment in a busy port.

3.3.2. Risk assessment methods used in risk-based design

In the previous section, we identified risk-based design (goal-based design oriented toward functional requirements of safety controls) as the most common approach among studies presenting practical design approaches to MASS systems. Considering the central role of risk-based design in MASS systems, we examine the studies contributing to this topic more closely. The result is a list of fifteen studies, among which we identified nine different risk assessment methods. These are listed in Table 5 along with their respective definitions and corresponding studies. Note that while each method describes a distinct modelling approach for risk assessment, they were not exclusive, often appearing in combination with other approaches.

3.3.3. Human-computer interaction paradigms

In this section, we examine distinct human–computer interaction (HCI) paradigms adopted in the reviewed studies. We observed that such paradigms yielded different hypotheses about MASS functionality and are therefore important elements when considering the breadth of relevant design approaches. Overall, we identified two broad HCI paradigms, which we define in Table 6 along with their corresponding articles. These paradigms were described in sixteen of the forty-two studies in the review.

Situation Awareness (SA) was the most common HCI paradigm, referred to in nine of the studies. Surprisingly, almost all studies adopting SA also adopted elements of sense-making in their treatment of interactions. Sense-making represents a distinct cognitive paradigm from SA in that it links cognitive processes to the body, something that SA does not do. This unlikely combination may in part be a result of MASS operators' working remotely. Operators' bodies were in the SCC while their actions involved distant MASSs, a cognitive discord that was bridged by adopting elements from both SA and sense-making paradigms.

Task Analysis (TA) was another common framework used to describe HCI. The task analyses presented in Ramos et al. (2019, 2020a, 2020b) provide detailed classification of SCC operators' roles and tasks, in their case for MASS without operators onboard. In Ramos et al. (2019), for example, the authors present a "task description" listing specific tasks goals and sub-goals, as well as descriptions of possible errors corresponding to the defined tasks (pp. 40–41). For example, the first task goal they list is "supervise safety status of the ships," which has the sub-task "monitor the screens" and the corresponding possible error "not checking information on the screens" (p. 40). This way of conceptualizing the MASS system was practical for modelling risks and for guiding interface design.

The study of human factors also provided a lens on HCI. For example, Yoshida et al. (2021) frames human–computer interactions as a source of stress for SCC operators. They argue that stress "might sometimes increase due to information overload by receiving enormous amounts of visual data to compensate for the lack of the feeling of the environment inside or outside a ship" (p. 2). In another study, Wróbel et al. (2021), studying the various human factors emerging from human-computer interaction within the MASS system, adopt a systematic method called "Human Factors Analysis and Classification System for Maritime Accidents (HFACS–MA)." Both studies indicate that human factors play an important role in design of safe interactions in the MASS system.

3.3.4. Empirical testing and prototyping

Fifteen studies presented results of empirical testing and observations about MASS systems. Considering MASS systems are still in a developmental phase, such results are valuable given that they can contribute to further design iterations and extended research. The empirical material came in a variety of forms, ranging from expert workshops and interviews to simulator tests and field trials with MASS prototypes. In Table 7 we list an overview of empirical testing and data collection efforts in the reviewed studies.

While the studies in Table 7 generated new observations to inform their work, several others used historical data. These other efforts, while not contributing new data to the research community, did present useful data collection methods. For example, Thieme and Utne (2017) and Fossen and Fossen (2018) both use field data collected from Trondheim Fjord in Norway, which is a designated a test area for autonomous ships. The former use field trial data from autonomous underwater vehicles (AUVs) to make inferences about safety performance indicators for MASS operations, while the latter use live Automatic Information System (AIS) data to test ship motion prediction and visualization algorithms intended specifically for remote operators. In another study, Wu et al. (2021a) collect data from the ship log of a research vessel transiting between two locations. Taking this information to represent successfully voyages, dynamic safety indices were generated based on the extent to which route deviations were detected in real time. Wu et al. (2021b) accomplish something similar for AIS traffic data to inform collision avoidance maneuvering at a VTS or SCC.

3.4. Given that automation is increasing in shipping, what research gaps, regulatory obstacles, and technical shortcomings represent the most significant barriers to real-world implementation?

In this section, we explore the extent to which the lack of full-scale autonomous ships currently implemented in the field is due to research gaps, regulatory obstacles, and technical shortcomings. Together, the reviewed studies described MASSs along a spectrum of technology readiness levels. Some authors argue that MASS are ready for implementation so long as the appropriate operational strategies are in place (e.g., Rødseth et al., 2021; Valdez Banda et al., 2019); others argue out that MASS do not currently exist in operational scale because of unresolved design obstacles (e.g., Kooij and Hekkenberg, 2021; Relling et al., 2021). Here, we investigate four relevant themes in the effort to identify and assess what barriers currently exist on the path from concept design to real-world implementation: (1) emerging risks in human-AI interaction, (2) regulatory obstacles, (3), training for SCC operations, and (4) infrastructure needs to support mixed conventional ship and MASS traffic. These categories shed light on the directions of current research in the field and on where more work is needed.

3.4.1. Emerging risks in human-AI interaction

The MASS system introduces novel challenges related to integrating human supervisory control within a highly automated system. In the reviewed studies, a new risk picture emerged characterized by uncertainties involving how human-AI interaction will be accomplished in real-world conditions. Here we examine how human-AI interaction was positioned as a central factor in twenty-two of the analyzed studies, underpinning risk assessment and design of safety controls.

In Section 3.3.2, we presented studies that applied risk assessment methods to study MASS systems. The result of such risk assessments typically materialized as a list of potential hazards, with the hazards near the top of the list contributing the largest share of overall risk. Chang et al. (2020), for example, review all published hazard identifications for MASS operations and evaluate the most salient hazard categories. Of these, they conclude that "interaction with manned vessels and detection of objects" represent the most significant contributors to overall risk (p. 10). Wróbel et al. (2018) (which was one of the studies Chang et al. reviewed), also determined that "interaction between shore-based facilities and legal or organizational [entities]" was a significant source of hazards. This finding is corroborated in Guo et al. (2021), in which the failure probability of an autonomous urban passenger ferry is found to be "most sensitive to the failure of the remote supervisor's intervention" (p. 7). Eriksen et al. (2021), taking a different approach, frame interactions in terms of preventative maintenance and conclude that so long as conventional ship machinery is used in MASS, maintenance requirements cannot be met without personnel present on board. The overall risk picture of the MASS system, in other words, remains unclear until we design its functional requirements around human interaction. Anticipating that risks arise from interactions within a socio-technical system, the "H-SIA" method was developed specifically in Ramos et al. (2020a) to model error propagations from human-AI interactions.

Among obstacles towards understanding the mechanisms behind human-AI interaction, lack of research and especially empirical data emerged most clearly. It was expressed in twenty-two of the forty-two articles analyzed. Of these, the most cited was related to lack of data about how failures propagate between interaction of hardware, software, and humans (Bolbot et al., 2021; Fan et al., 2021; Guo et al., 2021; Ramos et al., 2019, 2020a, 2020b; Thieme et al., 2018). Also wanting was research about risk management frameworks appropriate for MASS systems (Chang et al., 2020; Ramos et al., 2020a; Thieme et al., 2018; Thieme and Utne, 2017). Some studies also pointed out that observation-based data is needed about hazards for MASS in order to shed light on mitigation strategies in the aim of eventually being able to quantify their associated risk levels (Thieme et al., 2018; Valdez Banda et al., 2019; Wróbel et al., 2018). Similarly, several studies highlighted the lack of research about human error probability quantification for human-AI interaction especially for emergency response (Ramos et al., 2019; Zhang et al., 2020; Zhou et al., 2021). Other areas where lack of empirical data was reported as a barrier to real-world implementation of MASS applied to safety management (Størkersen, 2020), reliability of MASS machinery (Eriksen et al., 2021), human-AI system interface design (Fan et al., 2021; Huang et al., 2020), and competence requirements for SCC operators (Kim and Mallam, 2020; Sharma and Kim, 2021). Interaction between MASS and conventional vessels was also an area lacking data (Relling et al., 2021; Thieme et al., 2018), as were observations about how human factors affect SCC operations (Hannaford and Hassel, 2021; Kari and Steinert, 2021; Wróbel et al., 2021; Yoshida et al., 2021).

3.4.2. Regulatory obstacles

Gaps in regulations for unmanned ships were identified as a major barrier to MASS system implementation in the analyzed studies. They were prominently featured in eleven out of the forty-two papers analyzed. Common to all studies was the finding that conventions like COLREGS, SOLAS, and STCW, which are all regulated by IMO, are predicated on seafarers being onboard the ship. Legal precedents are also based on human crew onboard the ship. Katsivela (2020), a maritime law scholar, argues for amendments to existing conventions that would allow for MASS and SCC to be regulated as conventional ships. Although the author suggests that these definition-based amendments are relatively straightforward to implement, the need for them is described as "urgent," as are efforts towards harmonization and standardization of terms used across many different regulatory agencies at both international and national levels. Katsivela (2020) and Yoshida et al. (2020) both highlight Rule 5 in COLREGS as an example of how current regulations may have limited applications to MASS. The rule states that "Every vessel shall at all times maintain a proper look-out by sight and hearing as well as by all available means appropriate in the prevailing circumstances so as to make a full appraisal of the situation and of the risk of collision." Both authors suggest that while sensor fusion is showing potential in accurate detection and classification of objects, the appraisal of the "situation and the risk of collision" appears to necessitate the presence of a human. Wu et al. (2021b) argue that COLREGS is "human-centered" and "formulated in an intuitive manner," adding that its qualitative nature imposes challenges on the development and implementation of rule- or machine-learning-based approaches to collision avoidance. Katsivela is particularly attentive to the vocabulary used in maritime conventions and how it accentuates gaps between regular ships and MASS. For example, the wording in international conventions like STCW uses phrases like "a personnel onboard," "those on board," and "seafarers on board," which have been attributed to the obligation of the shipowner in maritime law to "provide for a seaworthy vessel" (United Nations, 1982). Also, the definition of "seaworthiness" has legal precedent in maritime law (Tetley, 2008) requiring that "The vessel be capable of withstanding the ordinary perils of the sea, be fit for the proposed trip and be crewed by a competent crew." Katsivela (2020) suggests clarifications of definitions for "ship," "navigation bridge," and "seaworthiness" as they are defined in the current conventions to include the possibility of unmanned ships with crew on land. Kooij and Hekkenberg (2021), in their proposed incremental crew reduction strategy for MASS systems Wróbel, 2021, also meet barriers for allowing seafarers to be responsible for ships at a distance. This regulatory change, they write, is the last step in the process of moving seafarers to land-based work like that envisioned at the SCC.

Conventions for seafarer training and competence requirements were also lacking prescriptions for meeting the novel demands of skilled SCC work. These conventions are laid out in STCW, which Kim and Mallam (2020) and Sharma and Kim (2021) investigate for applicability to MASS. Out of the sixty-six knowledge, understanding, and proficiencies (so-called "KUPS") listed (IMO, 2017, Table A-II/1), Sharma and Kim (2021) surmise that only twenty-six are relevant for MASS operations, while many whole new ones will be needed. Moreover, it was found that STCW leadership strategies need to be re-evaluated (Kim and Mallam, 2020) and that potential pitfalls like "over-reliance on sensors" warrant special attention (Hannaford and Hassel, 2021). Saha (2021) also finds that STCW presents a clear gap when it comes to SCC operators and suggests new competency requirements in the form of "system understanding, communicational and technical knowledge, and maritime competence" (p. 6).

Bačkalov (2020) and Nzengu et al. (2021) focus on regulations for inland MASS. Nzengu et al. (2021) point out that inland MASS do not fall under IMO jurisdiction, since "their operations and their context differ significantly from short-sea and ocean-going ship applications" (p. 359). The authors conclude that current regulations do not allow for autonomous MASS because they by definition rely on human functions onboard. Bačkalov (2020) presents detailed suggested amendments to updated regulatory framework for inland MASS, including introducing the notions of the "remote operator" and the "remote control center," as well as refining the "notion of the Master of an inland vessel" (p. 11).

Rødseth et al. (2021) focus on the approval process of MASS systems, comparing six design guidelines published by flag states and classification societies. The authors find that the six guidelines are largely in agreement that approval will hinge on submitting "descriptions of the MASS (system) design, [its] intended operations, and the environment it is intended to operate in." However, since it was not immediately clear from the guidelines how best to describe the system, the authors develop the concept of the "operational design envelop." Intended to help designers describe the MASS system for the approval process, the operational design envelop encourages description of how responsibilities between human and automation are shared, independent of project-specific factors like geographic location and operation.

3.4.3. *Training for SCC operators*

Eleven studies addressed training needs of MASS operators working at the Shore Control Center – so-called "SCC operators" (SCCOs). The knowledge gap in what training and competency requirements were needed for these operators was framed as a barrier to MASS implementation.

Four of the studies framed the need for training in terms of risk, where training was seen as a safety control. Wróbel et al. (2018) and Fan et al. (2020), for example, argue that insufficient training of SCC operators can lead to inadequate safety control functions. By this logic, the reverse also holds true: "…a well-trained and experienced operator having adequate control over the vessel can find the best solution to the situation encountered" (Wróbel et al., 2018, p. 342). Zhang et al. (2020) go one step further, presenting the effect of "insufficient training" in quantitative terms among other human error factors present in SCC work. Thieme et al. (2018), in assessing various risk modelling techniques for applicability to MASS systems, find that training, competence, and experience feature repeatedly as factors influencing risk.

The remaining seven studies focused on what type of competencies are needed at the SCC. Ahvenjärvi (2016) was among the first to point out that special training will be needed to meet the demands of interacting with manned and unmanned ships in the same area. Subsequent studies agreed that a lack of knowledge about how to train SCC operators is hindering MASS implementation. Katsivela (2020) express this problem from a regulatory perspective: "…existing regulations relating to the training and certification of the crew and, therefore, the provisions of the International Convention on Standards of Training, Certification and Watchkeeping for Seafarers 1978 (STCW) currently under the purview of the IMO's scoping exercise will have to be revised to accommodate the presence of a properly trained crew onshore" (p. 243). (Note that the scoping exercise is now completed and recommends further investigation into these competency needs, a topic we outlined in Section 1.3). Three studies compared competency requirements for conventional ships with those that are most likely required for MASS operations, orienting readers towards appropriate revisions to the STCW or to tailor-made SCC requirements (Kim and Mallam, 2020; Saha, 2021; Sharma and Kim, 2021). Baldauf et al. (2019), on the basis of results from a simulator experiment, argue that ship-handling skills and experience at sea are prerequisites for SCC operation. This is corroborated by Saha (2021) and Yoshida et al. (2020), who confer with subject matter experts and seafarers. Additional competency requirements also emerged, like those associated with digital communication and information transfer (Saha, 2021; Yoshida et al., 2020). The former concludes that, "the SCCO should possess a combination of maritime and technological competence to control the unmanned vessels with comprehensive knowledge of the remote vessel operational and monitoring system" (Saha, 2021, p. 11).

Sharma and Kim (2021) were also motivated to address the shortcomings in STCW to MASS operations, and systematically compare sixty-six KUPs set out in current convention (IMO, 2017, Table A-II/1). Their findings point to twenty-six KUPs that are no longer relevant and to emerging requirements for "technical and cognitive competencies." The former involves "IT skills, safety & security management skills, knowledge regarding engine room operations, electronic equipment and system integration" (p. 13), while the latter involves "non-routine problem solving, ability for self-regulation, critical thinking, mental readiness and systemic thinking" (p. 15). Sharma and Kim (2021), after accumulating survey results, report that most seafarers agree that training will change significantly with the onset of autonomous shipping and that skillsets will likely be similar to those of currently licensed seafarers.

The studies reviewed revealed that questions about training, until resolved, present a barrier to MASS implementation. This is the case whether framed as a safety control mechanism in risk science or as competency requirements human factors. Training, in this sense, encompasses not just skills, but qualifications, roles, and responsibilities that are at the core of safe MASS operations.

3.4.4. *Infrastructure to support mixed traffic*

Although autonomous vessels stand ready for implementation, the reality of marine traffic is that it will always be mixed, composed of both MASS and manned vessels sharing the water. This mixed traffic format will introduce unique challenges. Here we briefly summarize results from seven studies encompassing three relevant topics: (1) SCC coordination with VTS, (2) development of mixed traffic navigation aids, and (3) harmonization of rules and standards for aids to marine navigation.

It emerged from four studies that VTS will play an important role in coordinating MASS traffic. VTS is regulated globally by an organization called the International Association of Marine Aids to Navigation and Lighthouse Authorities (IALA). Katsivela (2020) refers to a 2019 report by IALA that affirms that the organization is preparing for the advent of MASS by monitoring technology development for MASS navigation services and by promoting harmonized rules and standards regarding data transferred via different services (IALA, 2019, pp. 10-11). Baldauf et al. (2019) focus on challenges associated with interaction between SCC and VTS. The authors' premise is that since that both infrastructures will have similar roles of coordinating marine traffic from a centralized control room, some degree of cooperation will be needed. This logic is also held by Relling et al. (2021), who examine how VTS procedures can be standardized to include MASS coordination. Their findings point to increased involvement: "VTS need to change their role from solving situations ad-hoc to assume a tactical responsibility in traffic planning and to resolve situations at an earlier stage" (p. 1). The authors underscore that this not does not replace SCC roles, but rather shares coordination tasks between SCC and VTS operators. Yoo and Lee (2021) were

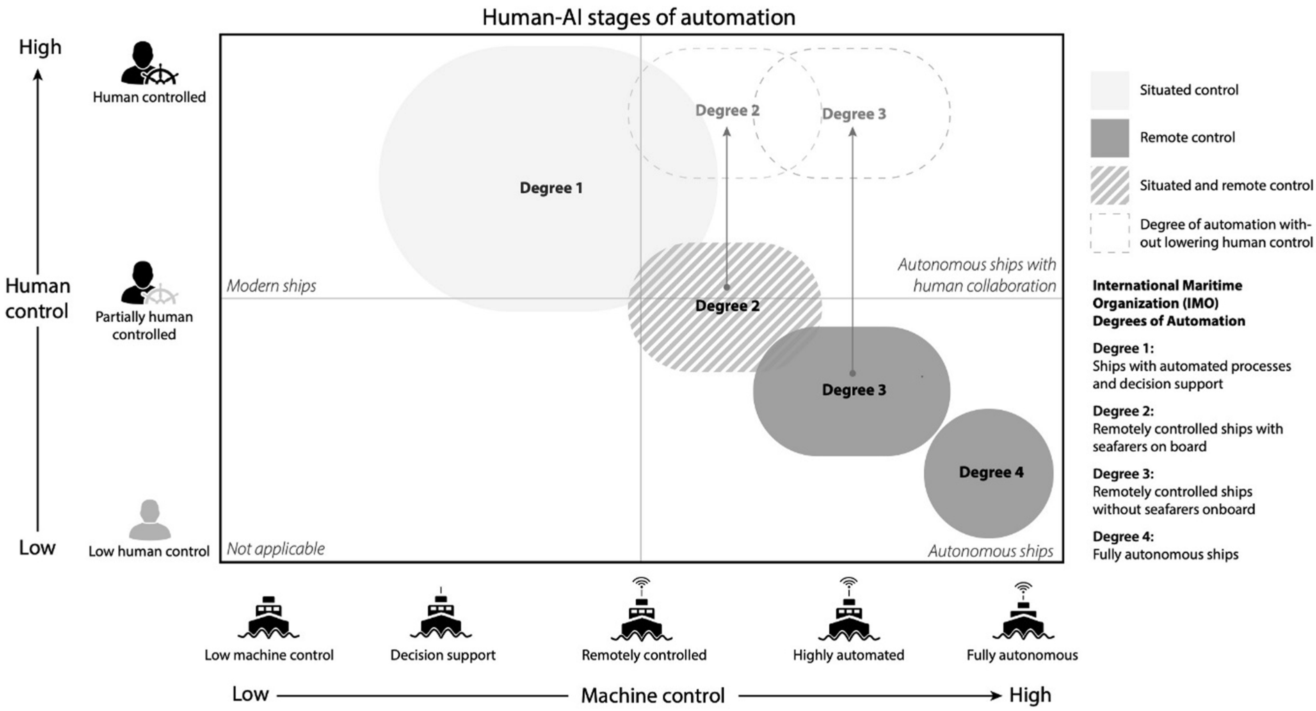

Fig. 5. IMO degrees of automation plotted onto the Human-Centered AI "Stages of Automation" framework adopted from Shneiderman (2020).

also motivated by what services VTS could provide in MASS traffic management. Their work came on the tail of IALA's announced intentions to work towards integrating "…autonomous vessels with conventional traffic, the information flow between MASS and shore authorities, and the related information exchange with conventional traffic" (IALA, 2019, p. 15). In this aim, the authors developed a "real-time collision risk assessment support system to improve the situational awareness of VTSOs [VTS operators] and MASS remote operators… regarding near-collision situations occurring in local waters" (p. 1).

An additional two studies argue that aids to navigation will have to be updated to support mixed traffic. Chang et al. (2020), after conducting a detailed risk assessment, found that "interaction with manned vessels" was among the highest contributors to overall risk of MASS operations. Thieme et al. (2018) also investigated mixed traffic from a risk science perspective, showing that traditional risk assessment methods that estimate collision frequency from conventional ship traffic will likely need to be updated for new traffic patterns that will inevitably emerge. The same authors highlight that navigation lights and buoys are used by deck officers and lookouts mainly as a type of visual ground-truth to radar and are in this way oriented uniquely to human navigation practices. Considering that MASS navigation will rely on the detection of such aids to navigation, consideration should therefore be directed towards making them more "visible" to sensor technologies.

Three studies specifically raised harmonization of rules and standards as a barrier to MASS implementation in mixed traffic. Katsivela (2020) defines harmonization as "any attempt, by whatever instrument, to minimize or eliminate discord between national commercial laws as they apply to international commercial transactions" (p. 241). Katsivela (2020) shows that the discord between MASS development, undertaken by a small number of institutions, are in discord with international standards for navigations infrastructure. The problem is especially evident for marine aids to navigation. These are regulated by IALA, whose initiatives to support digitalization of navigation (so called "e-navigation") still assumes that human operators are present on ships. Størkersen (2020) and Baldauf et al. (2019) similarly argue that harmonization of procedures across different procedures (e.g., MASS versus conventional ships, SCC versus VTS), will serve to address current obstacles to MASS implementation.

4. Discussion

In this section, we expand upon the themes uncovered in the systematic review, linking our original three research questions to farther-reaching discussions about how the introduction of AI is changing safety–critical work and about how risk scientists, designers, and researchers are addressing emerging challenges underpinning system design.

Sections 4.1 and 4.2 expand upon Research Question 1. Section 4.1 focuses on the need for "continuous monitoring and decision making support" made clear in the Results. Because this need is at odds with the traditional LoA approach, which implies a one-off focus on automation at the cost of human control, we explore alternatives to the traditional approach. Section 4.2 addresses the debate about "human error" in design, especially in the context of the de-facto risk baseline that came to light in the Results; namely, that MASS systems should be "at least as safe as" conventional ships. If automation is foremost to reduce human errors, should not MASS systems by this logic be significantly safer than conventional ships? We consider what new risks are introduced when increasing interactions with AI systems, as well as how these new risks may be addressed.

Sections 4.3–4.5 expand upon Research Question 2. Section 4.3 considers the popularity among the reviewed studies of cognitive approaches to human decision making modeling in the risk sciences. We consider limitations of such cognitive modeling approaches in light of contemporary research in the fields of computer supported collaborative work and neuroscience. Section 4.4 explores the viability of virtual simulation as a venue of systematic testing of MASS system, including a brief assessment of its strengths and weaknesses. Section 4.5 takes a second look at risk assessment methods for MASS systems. Considering the relatively wide spectrum of approaches identified in the Results, we raise the question: are some methods more appropriate that others?

Finally, Section 4.6 expands upon Research Question 3. Specifically, it addresses barriers to real-world implementation of MASS systems through the wide lens of safety management. The review results made clear that a special dexterity is needed to realize human-AI teaming flexible enough to deal with real-world demands. We examine the balance of operational flexibility with managerial control, drawing comparisons to established theories in organizational science.

4.1. The need to increase human control in navigation for AI applications

In Section 3.2.1 we reviewed forty-two studies about autonomous ships systems, guided by the question *what are the implications for human control?* Although it seemed like a paradox, the review showed that humans are needed to control autonomous ships. Part of the answer lay implicit in the way Level of Automation (LoA) taxonomies are used to describe human-AI collaboration. All taxonomies available (there were six in the studies reviewed) involved some degree of human control; no system is currently being designed for fully autonomous operations. Human control most often took the form of "Continuous monitoring and decision making support" (Section 3.2.2). Furthermore, the studies argued that this role ensured the safety of the system. Extending this idea here, we can interpret this as positioning operators as "active backup" to the automated system, where the operator is a safety control to the automated system. It seems counterintuitive, then, to categorize LoA by degree of autonomous control gained over human control lost, when in practice both are needed to ensure safety. For similar reasons, Eriksen et al. (2021) deliberately avoided using an LoA designation at all, noting they do not apply to how ship machinery systems are currently maintained. Rødseth et al. (2021) also concludes that none of the standard LoA systems do justice to explaining operators' time-critical role in the MASS system. This leaves the possibility open that other frameworks exist that may more appropriately handle the collaborative aspects of human and machine teaming as compared to the LoA frameworks adopted in this review. Here we explore alternatives to the traditional LoA approach.

One framework was recently proposed by Shneiderman (2020) who introduced a two-dimensional "Stages of Automation" in place of the more prevalent one-dimensional "Levels of Automation." This framework is what Shneiderman calls Human-Centered Artificial Intelligence (HSAI), which aims to improve AI-based systems design by jointly increasing human and machine control to achieve "Safe, Reliable, and Trustworthy" systems. Shneiderman's work focuses predominantly on the car industry, which is currently leading innovations in autonomous transportation. Specifically, the work is in response to current trends in autonomous car design that are leading to purportedly unacceptable risks by combining high autonomy with low human control. The current LoA framework used in the car industry is published by the Society of Automotive Engineers (SAE) and features six levels, ranging from Level 0 to Level 5 (SAE International, 2017). But like the frameworks designed for MASS operations encountered in the review, the SAE's focus on automation implies that human control decreases with each increment on the scale. Considering that the four-degree IMO LoA was the most common framework adopted in the reviewed studies, we mapped them onto the "Stages of Automation" framework for direct comparison (Fig. 5). IMO Degrees 1 and 2 fit roughly in the intersection of the human–machine control quadrants. Examples of these types can be seen in modern ships and early MASS designs today. IMO Degree 3 fits into the low human control-high machine control quadrant and is where most attention in MASS development is afforded today. It is surprising that the only element setting Degree 3 and Degree 2 ships apart is the complete absence of people on board, implying that MASS autonomy is improved simply by superseding human control. Finally, IMO Degree 4 appears to exist only for symbolic reasons, considering not one of the papers reviewed considers full machine autonomy technically feasible. This way of framing the popular IMO Degrees of Automation suggests that attention should be focused on "raising" Degree 2 and 3 from their current positions towards higher levels of human control, as illustrated by the arrows in Fig. 5. Fig. 5 also illustrates a trend tracing automation development in modern ships (upper left quadrant) towards the "fully autonomous" vision (lower right quadrant"). But considering the infeasibly of the fully autonomous model, one must ask why technology trends tend to orient towards such a model. The reviewed studies, after all, repeatedly emphasized the importance of improving human control, not supplanting it altogether with automation. Consider, for example, the results presented in Section 3.2.2, where we reported twenty-three studies specifically addressing humans' roles in autonomous ship systems. It follows that re-aligning design efforts toward the upper right quadrant, labelled "autonomous ships with human collaboration" in Fig. 5, may stand as a more appropriate guiding model. Of the reviewed studies, Rødseth et al. (2021) came closest to the "Stages of Automation" approach by avoiding adherence to current LoA taxonomies and instead presenting a two-dimensional grid with human control versus automation. "We believe that this is a useful approach," they write, "… as it clearly defines the responsibilities of the automation system versus humans…" (p. 5).

The review also led to a clearer understanding of specific tasks that necessitate shared human–machine control in safe MASS systems. For example, Huang et al. (2020) and Wu et al. (2021b) remind readers that interpreting regulations like COLREGS in rule-based programming code is challenging, suggesting that an AI-based collision avoidance system may require some level of collaboration with a human operator. Moreover, a recent empirical study by Rutledal et al. (2020) shows that collision avoidance is about more than just following codified rules. The study recorded instances in marine traffic where adherence to COLREGS was intentionally broken to resolve a give-way vessel in possible collision situations. Nonetheless, efforts in encoding COLREGs both in traditional symbolic AI and uncertainty-based or heuristic methods have seen considerable attention for over two decades (Statheros et al., 2008; Tam et al., 2009). Efforts in this area generally improve in accuracy over time, with recent work showing high accuracy in adopting Rapidly-Exploring Random Tree (RRT) algorithms (Chiang and Tapia, 2018; Zaccone et al., 2019). However, limitations like adverse weather conditions and non-AIS vessel targets suggest that more work is needed before these systems can be robust enough to make decisions on their own, relegating them for now to the lowest IMO Degree of "decision support" automation. The "Stages of Automation" framework may thus have a practical application for designing such collaborative collision avoidance systems, addressing the need to reconcile outputs of algorithms and the context of human control in varying conditions. Platooning, which combines human control and automation in an elegant way, represents one such reconciliation of human and machine control collaboration (e.g., Colling et al., 2021; Munim et al., 2021).

Aside from interpretation of COLREGs, several other important elements of seafaring appeared more straightforward for humans to accomplish than for AI systems. This included, for instance, managing Very High Frequency (VHF) marine radio communication and cargo loading and unloading. Checking the literature, we found that it has long been observed that tasks humans find naturally straightforward are often firmly outside the grasp of advanced AI. First articulated by Moravec (1988, pp. 15), this principle has since been coined the "Easy things are hard" paradox in AI. This has important implications especially for regulatory gaps that were identified in the review, because requirements of "seaworthiness" as well as statutory rules for "proper lookout by sight and hearing" (COLREGs) and responding to others in distress at sea (SOLAS) should be considered firmly within in the realm of human interpretation and control – firmly beyond the reach of AI methods.

4.2. New challenges emerging from increased human-AI interaction

Studies assessing risk of MASS operations often surmised that MASS should be "at least as safe as conventional ships" (Fan et al., 2020; Reddy et al., 2019; Relling et al., 2021; Saha, 2021; Thieme et al., 2018; Valdez

Banda et al., 2019; Zhou et al., 2021). Similarly, Utne et al. (2020) cite a DNV GL position paper recommending a goal-based Autonomous Ship Code (ASC) regulated by IMO stating that "autonomous and remote-controlled ships shall be as safe as conventional ships of the same type" (DNV GL, 2018). The question we raise here is whether this "at least as safe as" baseline is appropriate from a risk perspective. The predominant argument, after all, for introducing autonomous systems is that they will improve safety by reducing instances of "human errors" – sources widely held to account for 75–95% of marine casualties (Rothblum, 2000). Following this logic, should not system design produce *safer* systems – and by a significant margin? The automotive industry, to contrast with shipping, appears to follow the latter logic. Leading developers of autonomous cars proclaim significant improvements in safety when compared to their conventional counterparts (e.g., Tesla, 2019; Waymo, 2020). The logic follows from a "human error" narrative similar to the shipping case, with oft-cited statistics attributing attributed the bulk of accidents to human driver errors (e.g., U.S. National Highway Traffic Safety Administration, 2015, which sets the statistic at 92–96%). The comparison to automation in the car industry is not without precedent; Rødseth et al. (2021) adopted their central concept of "operational design domain" from design approaches of autonomous cars. Some important differences between cars and ships do exist, however, which the authors diligently highlight. "Ships," they write, "[and] especially large ones, have a much higher cost and damage potential compared to road vehicles… Låg, 2019 [also] most merchant ships move slowly and have more space to use for maneuvering" (p. 5). As a bottom line, though, safety baselines for MASS based on conventional ships should be brought to question. This is especially the case given that "human error" seems impossible to fully circumvent in ship systems, in that they will always exist by nature of human interaction, regardless of how much autonomous control is incorporated (Wróbel et al., 2017).

The review also showed that assessing the performance of autonomous systems within frameworks of conventional ship navigation occasionally led to experimental pitfalls. For example, in Man et al. (2018), the authors conducted an experiment in a full-mission bridge simulator set up such that participants operated vessels that appeared to be navigated autonomously when in fact they were be steered by researchers in another room. The authors set out to study the interface design of the SCC by measuring how they used bridge equipment to handle scenarios of remote operation. However, they concluded that "when the operators utilize commonly used navigational and collision avoidance technologies but in a different way, they would have problems in developing sufficient situation awareness for remote supervisory control tasks (pp. 241)." In other words, by attempting to study needs of remote operation by using tools found on a conventional ship's bridge, the researchers could only conclude that the tools on a conventional bridge were not sufficient. A similar problem was described by Baldauf et al. (2019), who, after running experiments with participants in a ship simulator to investigate interactions between a simulated MASS and VTS, observed that the results were limited by the equipment available. While concluding that VHF marine radio was needed and that frequent cross-checks between bridge navigation equipment and the bridge windows were important, the results generally failed to link results to the underlying research question about MASS interaction design. The primary conclusion that one can draw from the two studies is that conventional navigation equipment is probably not sufficient for novel MASS applications. They serve as examples of what can happen when making inferences about paradigm-challenging observations in this way, analogous to over-fitting a model to observations, leading to results that do not satisfy hypothesis testing. The pitfall is forewarned by Kari and Steinert (2021), who write, "the working environment in the SCC is completely different from the traditional onboard bridge" (p. 17).

Increased interaction with AI systems appears to come with increased risk. Zhou et al. (2021) make the case that with every new interaction comes a security vulnerability, increasing the risk of cyber-threats. Meanwhile, Yoshida et al. (2021) argue that increased AI interaction results in higher mental workload among seafarers, subsequently undermining their performance. The latter also show that navigators use lookouts at their "eyes," trusting them differently than they would sensors in an automated lookout system. "Human–human communication is mentally more comfortable," they write, "than using autonomous support and navigation system in this situation" (p. 19). The findings are in line with the so-called "cooperative eye hypothesis," which explains that humans evolved to have large sclera (whites of the eyes) to be able to follow the gaze of others in cooperative activities relying on joint attention and communicative interaction (Kobayashi and Kohshima, 2001, 1997; Tomasello et al., 2007).

In a similar line of thinking, the increased human-AI interaction presented by the MASS system was framed by Sharma and Kim (2021) as "increasing invisible interactions." The authors continue: "It is paramount for the crews onboard to be able to have a holistic and systemic understanding of the systems and its interactions, to be able to comprehend the complexity, to evaluate the interrelations of subsystems and to subsequently generate the best decisions and course of actions" (p. 17). Whether it was framed as misconstrued safety statistics, vulnerabilities to cyber-attack, or compromised human factors effects on work performance, it appeared that increased interaction between human work tasks and AI functions contributed to a significant extent of the added risk in MASS systems.

4.3. Limits of cognitive models of situation awareness for risk modeling

In Section 3.3.2 we showed that state-of-the-art risk assessment methods of MASS operations included the human-in-the-loop by adopting simplified models for cognitive processes for human operators. In particular, the IDAC model described by Chang and Mosleh (2007) played an important role in H-SIA. In Ramos et al. (2019), the authors identified cognitive errors stemming from tasks using the IDAC framework by re-describing task goals (in terms of sub-goals and high-level goal plans) until one element in the cognitive model (Information, Decision, Action) could be isolated and subsequently linked to errors within a network of possible events, forming a coherent risk picture for the system. This approach is exemplary of human cognition modeling, which is useful in risk modeling by virtue of its predictive power of human behavior in novel situations. However, limitations of modelling cognitive processes of operators in control applications have been a source of debate in the literature (Endsley, 2015; Flach, 2015). For at least four decades, neuroscientists have accepted that modeling human cognition is faced with the apparently insuperable challenge of infinitely many mechanisms that can generate any given observation (Anderson, 1978 Shneiderman, 2016). Recent theories in neuroscience are showing promising ways to reconcile constraints of cognitive models with peoples' underlying irrationality, including for decision making (Lieder and Griffiths, 2020). Their applicability to risk modeling, however, has not yet been explored. There also remains important social elements to cognition, which have been investigated in depth in the case of bridge navigation by Hutchins (1995). Heath and Luff (1991) also demonstrated, with the case of control room operators for the London Underground, that "awareness" existed in a collaborative sense, implying that decision making emerges in a team rather than individual context for control room work. Limitations in modelling human cognition for use in risk assessments of human-AI interactions for collaborative control room work remain under-explored.

4.4. Virtual simulation methods

In Section 3.3.2 we uncovered a dilemma particular to the problem of risk modeling for new applications like MASS operations. This inherent paradox is as follows: risk models, which are needed to predict risk in novel operations, rely on historical data that do not exist because the operations in question are novel. While risk identification methods like STPA and expert judgment-based techniques like Bayesian Networks (BNs) circumvent the need for large amounts of historical data, verification will inevitably depend on empirical observations. The strategy of building test prototypes and operating them in controlled areas sanctioned for autonomous vessels is one strategy already adopted by some contemporary research efforts. But this approach is also expensive and, as observed by researchers like Baldauf et al. (2019), Peeters et al. (2020a), and Thieme et al. (2018), is faced with uncertainty symptomatic of early design phases. One potential solution lies in simulation. Several studies we analyzed highlighted the important role simulators play in investigating emerging human-AI interaction challenges. For example, Kari and Steinert (2021) write that simulator experiments will "help to assess and evaluate the role of human–machine and human–human interactions" (p. 17). Saha (2021) also points to benefits of simulators for training and assessment of SCC operators. In simulation, virtual versions of sensors sense a virtual world, allowing researchers and developers to study how the AI works without resource-intensive field testing. Scenarios can be built and tested rapidly, including emergency scenarios that are impractical or impossible to coordinate in real life. Graphics can make the virtual worlds immersive, opening possibilities to training and testing human factors and interaction solutions. The automotive industry has adopted virtual simulation to research autonomous cars, with one notable example being the open-source platform CARLA (Dosovitskiy et al., 2017). Recent efforts aimed at developing analogous platforms for MASS systems have emerged (Vasstein et al., 2020). Given that risk scientists broadly recognize the potential value of virtual simulation, and given its successful applications in autonomous car research, we can expect that future work will rely upon simulation before operation of MASS is a reality. Still, simulation is not without its limitations. The method depends upon accurate scenario development and is faced with the so-called "long tail problem" that precludes the simulation of rare events encountered only in the wide spectrum of possible events characteristic of the real world. Also, introducing MASSs will inevitably change the behavior of conventional marine traffic once implemented, a second-order effect that that alludes simulation.

4.5. Risk assessment methods for MASS systems

Risk assessment proved to be a popular topic among studies contributing to the theme of MASS systems. In Section 3.3.2, we analyzed fifteen studies that contributed risk assessments of MASS systems, which in total adopted nine different methods. The breadth of risk assessment methods was high compared to the number of studies adopting them. This raised the question of whether some methods were more suitable than others for MASS applications. This question was investigated in Thieme et al. (2018), who pegged System-Theoretic Process Analysis (STPA) as among most promising risk assessment methods for MASS system hazard analysis. Four of the reviewed studies used STPA for this purpose (Utne et al., 2020; Valdez Banda et al., 2019; Wróbel et al., 2018; Zhou et al., 2021), making it the most popular technique along with BNs. One reason for its apparent appeal is that STPA does not rely upon large amounts of historical safety records for purposes of inference. This is a virtue shared by BN techniques, which, as expressed by Thieme et al. (2018), offer "flexibility of modeling methods and input from experts" (p. 151). Four studies demonstrated, with detailed case studies, how BNs could be implemented into risk models (Chang et al., 2020; Guo et al., 2021; Utne et al., 2020; Zhang et al., 2020). BN approaches framed input from experts as a valuable source in the absence of operational data; this was in contrast to other studies whose authors, while also relying on expert judgment, framed qualitative data as an input source with limited value in the absence of objective, statistical data (e.g., in the FMEA technique adopted by Fan et al., 2021 and FSA approach in Bolbot et al., 2021).

Examining STPA approaches more closely, Wróbel et al. (2020), who champion the technique, demonstrate that organizational and human interaction issues remain under-represented in risk assessment, while focus on the technological components of the system receive the locus of attention. Wrobel et al. (2021) argue that it may even be impossible to quantify human errors in shipping applications. STPA circumvents the needs to set percentages to human errors, and instead focuses on how human interact with a system in a real-world setting. "Why bother quantifying such a complex social phenomenon [as human error]," ask the authors rhetorically, "using various and incomparable approaches, instead of investigating how to fit the human element into a technical system so that both operate at their optimum?" (p. 10). STPA is not without its critics. Bolbot et al. (2021), for example, highlight that STPA does not offer practical recommendations for completing a hazard identification. The authors opted instead for a traditional Formal Safety Assessment (FSA) approach, which they claim is also more straightforward to approve. The first risk assessment published on MASS also used FSA, highlighting similar reasons for doing so (Rødseth and Burmeister, 2015).

Other risk assessment methods framed safety controls in terms of preventative maintenance (Eriksen et al., 2021) or in terms of safety indices based on deviations of MASSs traversing known routes (Wu et al., 2021a). Like STPA and BNs, these methods, too, rely upon expert judgment as a model input. Preventative maintenance requires the expert machinist for input; building safety indices from navigation patterns relies on training Machine Learning algorithms on "expert" data. Even the FSA approach described by Bolbot et al. (2021) relied upon expert advice for hazard identification, which is explicitly recommended in its guidelines for novel or innovative designs (IMO, 2013, pp. 10-12). In this way, all risk assessment methods in the review made use of a combination of qualitative data in the form of expert judgment and quantitative methods in the form of statistical inference.

Aside from STPA and BN methods, another risk assessment approach that showed promise was the Hybrid Causal Logic (HCL) model. Originally developed for the offshore petroleum industry by Røed et al. (2009), HCL describes event propagations and their consequences. Outcomes are then linked to BNs, allowing for probabilistic input of subjective aspects like human and organization behavior. The H-SIA method introduced by Ramos et al. (2020a) was inspired by the HCL approach but lacked the implementation of BNs. A full HCL modeling approach for MASS operations was demonstrated by Wu et al. (2020), but it focused on interactions between manned and unmanned vessels rather than on interactions of land-based operators and the unmanned vessels under their control. An HCL model for a MASS control system can lead to risk assessments that are updated based on operators' experience and continually revised over time.

Among other potentially useful approaches is the Functional Resonance Analysis Method (FRAM), which has been adopted in dynamic safety management in complex socio-technical systems like healthcare and aviation. It was assessed in Thieme et al. (2018) for MASS systems and found to hold promise. However, there are currently no FRAM studies in the literature applied to MASS systems.

4.6. Safety management holds clues for reconciling human and machine autonomies

Safety management will play a crucial role in human-AI systems for safety critical applications. Theories that underpin safety management in terms of safety and (human) autonomy may also help to frame the emerging problem reconciling safety and (machine) autonomy in terms of practical organizational mechanisms. This applies to MASS operations but also all transportation systems introducing human-AI teaming.

In this review, there was only one paper that directly addressed safety management for MASS operations. The author frames safety management as both a "source of engagement and frustration among seafarers" and suggests that remotely controlled vessels open the door to amending conventions set out in International Safety Management (ISM) Code specifically for increasing task flexibility for operators and for reducing cumbersome paperwork (Størkersen, 2020). Historically, research in work design has focused on striking a balance between human autonomy of front-line workers and centralized managerial control. In situations that demand dexterity from operators in the face of uncertainty, this balance means combining a high degree of both operational flexibility and organizational stability. The theoretical groundwork was established by Weick (1976) who framed this balance as a contradiction between centralization and decentralization and as a contradiction between safety and worker autonomy. Still, an analogy may be made between front-line human workers and the AI tools being currently developed. Drawing this parallel, we can extend the safety management frameworks to consider safety management for human-AI systems. Weick also theorized that specific mechanisms were present: so-called “loose couplings” that bridge the divide between safety (centralized, stable, managerial control) and autonomy (decentralized, flexible, task-based work). Recent work by Grote (2020) has picked up the debate, suggesting three coupling mechanisms that can be used in safety management towards High Reliability Organizations (HROs): safety rules, safety leadership, and safety culture. While specific examples like participatory rulemaking, transformational leadership, and safety climate outline practical tools in this framework, the work remains theoretical and without direct empirical verification. Regardless, many elements of the conventional work design debate can carry over to the debate in human-AI control. We outline three carry-over elements here. The first is that many aspects of personal worker safety are now moot considering that AI tasks can replace them. For example, SCC workers are not spending shifts at sea, virtually eliminating “slips, trips, falls, and vehicles accidents” traditionally associated with personal or occupational safety. New occupational hazards will present themselves at the SCC, but the hazards will be fewer and less consequential. The second is that principles we have drawn from studying HROs match the aims of human-AI systems in that the focus is on things going right, rather than counteracting things going wrong. For example, Rasmussen says, “Rather than striving to control behavior by *fighting deviations* from a particular pre-planned path, the focus should be on the control of behavior by *making the boundaries explicit and known* and by giving opportunities to develop *coping skills at boundaries* (Rasmussen, 1997, pp. 191). Rasmussen was referring to boundaries between worker control and managerial control, but they apply also to AI control and supervisory control. The idea of making “boundaries explicit and clear” echo the need to understand the AI system limitations as well as the underlying motivation of “Explainable AI” (Gunning and Aha, 2019). The “coping skills” at boundaries refer to the adaptability and contextuality of working in shared control between an AI system and human operators. The third and final element that we discuss here is the opportunity to learn from HROs for designing resilient human-AI systems. This is challenging, given the lack of empirical data on the subject. However, adopting Grote’s framework and the concept of mechanisms for loose coupling between human autonomy and safety, one could imagine how strategies from HROs can carry over to loose couplings between AI and safety. Examples include participatory rulemaking, whereby the operator provides regular feedback to train AI systems, and strong cultural norms supporting switches between modes of operation.

5. Conclusion

In this article, we summarized the research state-of-the-art for autonomous ship systems, surveying a growing body of scientific literature published in peer-reviewed journals encompassing risk and safety science, human factors and training, marine policy, and ocean and reliability engineering. Given that automation is increasing in shipping, we posed three research questions: (1) how is human control currently being adopted in autonomous ship systems? (2) what methods, approaches, and theories are being used to address safety concerns and design challenges? and (3) what research gaps, regulatory obstacles, and technical shortcomings represent the most significant barriers to their real-world implementation?

In a systematic review, we collated findings from forty-two relevant peer-reviewed articles. Although the review spanned five broad disciplines (marine policy, ocean engineering, human factors, reliability engineering, and risk science), two underlying themes underpinned all findings: safety and control. The two themes were closely linked, with safety framed as a property emerging from reconciling control between machine autonomy and human autonomy. From this perspective, machine autonomy executes defined tasks while human autonomy manages supervisory goals under changing external factors and uncertainty.

Starting with the first research question, we found that human control plays a critical role even for the most advanced autonomous ship systems. Autonomous ships and their land-based counterparts, so-called Shore Control Centers (also called Remote Control Centers or Remote Operating Centers), represent a collaborative system. Most researchers envision human roles that go above and beyond simple backup to the automation, involving roles in all operational phases (e.g., watchkeeping, preventative maintenance, cargo loading and unloading, and emergency handling). During regular navigation, the AI system should enhance operators’ decision making for route planning and collision avoidance, not replace it outright. In line with this perspective, new thinking is emerging about how to frame autonomous system design as advancing in stages of human-AI interaction, rather than in levels of machine control.

In line with the second research question, we synthesized current thinking about which risk assessment methods are best suited for autonomous ship systems. Our findings pointed to System-Theoretic Process Analysis (STPA) and Bayesian Networks (BNs) as the most appropriate tools to use in goal-based design of safety controls. These techniques do not rely on historical data about accident frequency, and instead leverage expert human input. A parallel can be drawn in this regard to Machine Learning algorithms and resilience thinking, two relevant fields that also frame human expertise as a source of power, not as a source of error. Current thinking raises safety management frameworks originally intended to bridge the divide between worker autonomy and managerial control as relevant for human-AI systems. Parallels were drawn in this sense to control coupling mechanisms bridging the divide between machine autonomy and human control.

Finally, addressing the third research question, we sampled the research state-of-art for the most significant obstacles standing in the way of real-world implementation off MASS systems. Notable obstacles included regulations, which hinge upon definitions ships as having seafarers physically onboard, and training requirements, for which gaps abound in current maritime conventions. The most significant obstacle involved the uncertainly surrounding human-AI interactions. Research is needed to shed light about how people work when teamed with AI, especially in safety–critical contexts.

Autonomous ships will soon sail in our oceans, coastal waters, and inland waterways. Current research shows that humans will have just as important a role in ensuring safety of autonomous ships as for conventional ships. The question remains how this role will be most appropriately carried out and best to coordinate it with AI system functions.

CRediT authorship contribution statement

Erik Veitch: Investigation, Methodology, Writing – original draft, Writing – review & editing. **Ole Andreas Alsos:** Funding acquisition, Project administration, Resources, Supervision, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors gratefully acknowledge the financial support provided by the Research Council of Norway under the MAROFF-2 project "Land-based Operation of Autonomous Ships" (LOAS) (grant number 296527) directed by project leader Kongsberg Maritime AS and partnered with the Institute for Energy Technology (IFE). The authors also acknowledge the support of the Department of Design at NTNU.

Appendix A:. Dataset with ID numbers

ID	Author, year	Title	Journal	Category
1	Man et al. (2018)	Human factor issues during remote ship monitoring tasks: An ecological lesson for system design in a distributed context	International Journal of Industrial Ergonomics	Human Factors & Training
2	Kim and Mallam (2020)	A Delphi-AHP study on STCW leadership competence in the age of autonomous maritime operations	WMU Journal of Maritime Affairs	Human Factors & Training
3	Ahvenjärvi (2016)	The human element and autonomous ships	International Journal on Marine Navigation and Safety of Sea Transportation	Human Factors & Training
4	Sharma and Kim (2021)	Exploring technical and non-technical competencies of navigators for autonomous shipping	Maritime Policy & Management	Human Factors & Training
5	Hannaford and Van Hassel (2021)	Risks and benefits of crew reduction and or removal with increased automation on the ship operator	Applied Sciences	Human Factors & Training
6	Kari and Steinert (2021)	Human factor issues in remote ship operations: Lesson learned by studying different domains	Journal of Marine Science and Technology	Human Factors & Training
7	Yoshida et al. (2021)	Identification of the relationship between MASS and the operators' mental workload	Journal of Marine Science and Technology	Human Factors & Training
8	Saha (2021)	Mapping competence requirements for future shore control center operators	Maritime Policy & Management	Human Factors & Training
9	Wróbel et al. (2021)	On the influence of human factors on safety of remotely controlled merchant vessels	Applied Sciences	Human Factors & Training
10	Baldauf et al. (2019)	Merging conventionally navigating ships and MASS: Merging VTS, FOC and SCC	International Journal on Marine Navigation and Safety of Sea Transportation	Marine Policy, Regulation & Navigation
11	Yoshida et al. (2020)	Regulatory requirements on the competence of remote operator in maritime autonomous surface ship	Applied Sciences	Marine Policy, Regulation & Navigation
12	Størkersen (2020)	Safety management in remotely controlled vessel operations	Marine Policy	Marine Policy, Regulation & Navigation
13	Katsivela (2020)	Unmanned vessels and regulatory concerns	The Journal of International Maritime Law	Marine Policy, Regulation & Navigation
14	Relling et al. (2021)	The contribution of Vessel Traffic Services to safe coexistence between automated and convention	Maritime Policy & Management	Marine Policy, Regulation & Navigation
15	Kooij and Hekkenberg (2021)	Identification of a task-based implementation path for unmanned autonomous ships	Maritime Policy & Management	Marine Policy, Regulation & Navigation
16	Nzengu et al. (2021)	Regulatory framework analysis for the unmanned inland vessel	WMU Journal of Maritime Affairs	Marine Policy, Regulation & Navigation
17	Peeters et al. (2020a, 2020b)	An inland shore control centre for monitoring or controlling unmanned inland cargo vessels	Journal of Marine Science and Engineering	Ocean Engineering
18	Fossen and Fossen (2018)	eXogenous Kalman Filter for visualization and motion prediction of ships using live AIS data	Modeling, Identification and Control	Ocean Engineering
19	Peeters et al. (2020a, 2020b)	An unmanned inland cargo vessel: Design, build, and experiments	Ocean Engineering	Ocean Engineering
20	Wu et al. (2021)	Sailing status recognition to enhance safety awareness and path routing for a commuter ferry	Ships and Offshore Structures	Ocean Engineering
21	Rødseth et al. (2021)	Towards approval of autonomous ship systems by their operational envelope	Journal of Marine Science and Technology	Ocean Engineering
22	Yoo and Lee (2021)	Collision risk assessment support system for MASS, RO and VTSO support	Journal of Marine Science and Technology	Ocean Engineering
23	Wu et al. (2021)	Navigating patterns analysis for on-board guidance support in crossing collision avoidance operations	IEEE Intelligent Transportation Systems Magazine	Ocean Engineering
24	Ramos et al. (2020a, 2020b)	Human-system concurrent task analysis for Maritime Autonomous Surface Ship operation and safety	Reliability Engineering and System Safety	Reliability Engineering
25	Thieme and Utne (2017)	Safety performance monitoring of autonomous marine systems	Reliability Engineering and System Safety	Reliability Engineering
26	Wróbel et al. (2018)	System-theoretic approach to safety of remotely-controlled merchant vessel	Ocean Engineering	Reliability Engineering
27	Utne et al. (2020)	Towards supervisory risk control of autonomous ships	Reliability Engineering and System Safety	Reliability Engineering
28	Reddy et al. (2019)	Zero-emission autonomous ferries for urban water transport: cheaper, cleaner alternative to bridges and manned vessels	IEEE Electrification Magazine	Reliability Engineering
29	Huang et al. (2020)	A ship collision avoidance system for human–machine cooperation during collision avoidance	Ocean Engineering	Reliability Engineering

(*continued on next page*)

(*continued*)

ID	Author, year	Title	Journal	Category
30	Eriksen et al. (2021)	An RCM approach for assessing reliability challenges and maintenance needs of unmanned cargo ships	Reliability Engineering and System Safety	Reliability Engineering
31	Valdez Banda et al. (2021)	A systemic hazard analysis and management process for the concept design phase of an autonomous vessel	Reliability Engineering and System Safety	Reliability Engineering
32	Bačkalov (2020)	Safety of autonomous inland vessels: An analysis of regulatory barriers in the present technical standards in Europe	Safety Science	Safety & Risk
33	Fan et al. (2020)	A framework to identify factors influencing navigational risk for Maritime Autonomous Surface Ships	Ocean Engineering	Safety & Risk
34	Ramos et al. (2020a, 2020b)	A generic approach to analysing failures in human-system interaction in autonomy	Safety Science	Safety & Risk
35	Zhang et al. (2020)	A probabilistic model of human error assessment for autonomous cargo ships focusing on human–autonomy collaboration	Safety Science	Safety & Risk
36	Thieme et al. (2018)	Assessing ship risk model applicability to Marine Autonomous Surface Ships	Ocean Engineering	Safety & Risk
37	Ramos et al. (2019)	Collision avoidance on maritime autonomous surface ships: Operators' tasks and human failure events	Safety Science	Safety & Risk
38	Chang et al. (2020)	Risk assessment of the operations of maritime autonomous surface ships	Reliability Engineering and System Safety	Safety & Risk
39	Zhou et al. (2021)	A system-theoretic approach to safety and security co-analysis of autonomous ships	Ocean Engineering	Safety & Risk
40	Bolbot et al. (2021)	A novel risk assessment process-application to an autonomous inland waterways ship	Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability	Safety & Risk
41	Guo et al. (2021)	Risk assessment of collisions of an autonomous passenger ferry	Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability	Safety & Risk
42	Fan et al. (2021)	Towards a Framework of Operational-Risk Assessment for a Maritime Autonomous Surface Ship	Energies	Safety & Risk

References

Ahvenjärvi, S., 2016. The human element and autonomous ships. TransNav: Int. J. Marine Navigation Saf. Sea Transp. 10.

Allianz, 2020. Safety and Shipping Review 2020: An annual review of trends and developments in shipping losses and safety. Allianz Global Corporate & Specialty, Munich, Germany.

Anderson, J.R., 1978. Arguments concerning representations for mental imagery. Psychol. Rev. 85, 249–277. https://doi.org/10.1037/0033-295X.85.4.249.

Bačkalov, I., 2020. Safety of autonomous inland vessels: An analysis of regulatory barriers in the present technical standards in Europe. Saf. Sci. 128, 104763 https://doi.org/10.1016/j.ssci.2020.104763.

Baldauf, M., Fischer, S., Kitada, M., Mehdi, R., Al-Quhali, M., Fiorini, M., 2019. Merging conventionally navigating ships and MASS-Merging VTS, FOC and SCC? TransNav: Int. J. Marine Navigation Saf. Sea Transp. 13.

Bolbot, V., Theotokatos, G., Andreas Wennersberg, L., Faivre, J., Vassalos, D., Boulougouris, E., Jan Rødseth, Ø., Andersen, P., Pauwelyn, A.-S., Van Coillie, A., 2021. A novel risk assessment process: Application to an autonomous inland waterways ship. Proc. Institution Mech. Eng., Part O: J. Risk Reliab. 1748006X211051829.

Burmeister, H.-C., Bruhn, W., Rødseth, Ø.J., Porathe, T., 2014. Autonomous unmanned merchant vessel and its contribution towards the e-Navigation implementation: The MUNIN perspective. Int. J. e-Navigation Maritime Economy 1, 1–13.

CCNR, 2021. Central Commission for the Navigation of the Rhine - Automation [WWW Document]. URL https://www.ccr-zkr.org/12050000-en.html (accessed 11.22.21).

Chang, C.-H., Kontovas, C., Yu, Q., Yang, Z., 2020. Risk assessment of the operations of maritime autonomous surface ships. Reliab. Eng. Syst. Saf. 207, 107324.

Chang, Y.H.J., Mosleh, A., 2007. Cognitive modeling and dynamic probabilistic simulation of operating crew response to complex system accidents: Part 1: Overview of the IDAC Model. Reliab. Eng. Syst. Saf. 92, 997–1013. https://doi.org/10.1016/j.ress.2006.05.014.

Chen, Z., Chen, D., Zhang, Y., Cheng, X., Zhang, M., Wu, C., 2020. Deep learning for autonomous ship-oriented small ship detection. Saf. Sci. 130, 104812 https://doi.org/10.1016/j.ssci.2020.104812.

Cheng, Y., Zhang, W., 2018. Concise deep reinforcement learning obstacle avoidance for underactuated unmanned marine vessels. Neurocomputing 272, 63–73. https://doi.org/10.1016/j.neucom.2017.06.066.

Chung, M., Kim, S., Lee, K., Shin, D.H., 2020. Detection of damaged mooring line based on deep neural networks. Ocean Eng. 209, 107522 https://doi.org/10.1016/j.oceaneng.2020.107522.

Colling, A., Hekkenberg, R., van Hassel, E., 2021. A Viability Study of Waterborne Platooning on the Lower Rhine. Eur. J. Transp. Infrastruct. Res. 21, 71–94.

Dallolio, A., Agdal, B., Zolich, A., Alfredsen, J.A., Johansen, T.A., 2019. Long-Endurance Green Energy Autonomous Surface Vehicle Control Architecture, in: OCEANS 2019 MTS/IEEE SEATTLE. pp. 1–10. https://doi.org/10.23919/OCEANS40490.2019.8962768.

Diaper, D., 2004. Understanding Task Analysis for Human-Computer Interaction. In: Diaper, D., Stanton, N.A. (Eds.) The Handbook of Task Analysis for Human-Computer Interaction, LEA Publishers, Mahwah, NJ.

Gl, D.N.V., 2018. Remote-controlled and autonomous ships in the maritime industry (Position Paper). Hamburg, Germany.

Dosovitskiy, A., Ros, G., Codevilla, F., Lopez, A., Koltun, V., 2017. CARLA: An open urban driving simulator, in: Proceedings of Machine Learning Research. In: Presented at the Proceedings of the 1st Annual Conference on Robot Learning, pp. 1–16.

Dunbabin, M., Grinham, A., Udy, J., 2009. In: An Autonomous Surface Vehicle for Water Quality Monitoring. Australian Robotics and Automation Association, Sydney, Australia, pp. 1–6.

Endsley, M.R., 2015. Situation awareness misconceptions and misunderstandings. J. Cogn. Eng. Decision Making 9, 4–32.

Endsley, M.R., 1995. Toward a theory of situation awareness in dynamic systems. Hum. Factors 37, 32–64.

Eriksen, S., Utne, I.B., Lützen, M., 2021. An RCM approach for assessing reliability challenges and maintenance needs of unmanned cargo ships. Reliab. Eng. Syst. Saf. 210, 107550 https://doi.org/10.1016/j.ress.2021.107550.

Evans, J.H., 1959. Basic design concepts. J. Am. Soc. Naval Engineers 71, 671–678. https://doi.org/10.1111/j.1559-3584.1959.tb01836.x.

Eykholt, K., Evtimov, I., Fernandes, E., Li, B., Rahmati, A., Xiao, C., Prakash, A., Kohno, T., Song, D., 2018. Robust physical-world attacks on deep learning visual classification. In: Presented at the Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 1625–1634.

Fan, C., Montewka, J., Zhang, D., 2021. Towards a Framework of Operational-Risk Assessment for a Maritime Autonomous Surface Ship. Energies 14, 3879.

Fan, C., Wróbel, K., Montewka, J., Gil, M., Wan, C., Zhang, D., 2020. A framework to identify factors influencing navigational risk for Maritime Autonomous Surface Ships. Ocean Eng. 202, 107188.

Flach, J.M., 2015. Situation awareness: Context matters! A commentary on Endsley. J. Cogn. Eng. Decision Making 9, 59–72.

Grote, G., 2020. Safety and autonomy: A contradiction forever? Saf. Sci. 127, 104709 https://doi.org/10.1016/j.ssci.2020.104709.

Grote, G., Weyer, J., Stanton, N.A., 2014. Beyond human-centred automation – concepts for human–machine interaction in multi-layered networks. Ergonomics 57, 289–294. https://doi.org/10.1080/00140139.2014.890748.

Gunning, D., Aha, D., 2019. DARPA's Explainable Artificial Intelligence (XAI) Program. AIMag 40, 44–58. https://doi.org/10.1609/aimag.v40i2.2850.

Guo, C., Haugen, S., Utne, I.B., 2021. Risk assessment of collisions of an autonomous passenger ferry. Proc. Institution Mech. Engineers, Part O: J. Risk Reliability, 1748006X211050714.

Chiang, H.L., Tapia, L., 2018. COLREG-RRT: An RRT-Based COLREGS-Compliant Motion Planner for Surface Vehicle Navigation. IEEE Rob. Autom. Lett. 3, 2024–2031. https://doi.org/10.1109/LRA.2018.2801881.

Hannaford, E., Hassel, E.V., 2021. Risks and Benefits of Crew Reduction and/or Removal with Increased Automation on the Ship Operator: A Licensed Deck Officer's Perspective. Appl. Sci. 11, 3569.

Hansen, H.L., Nielsen, D., Frydenberg, M., 2002. Occupational accidents aboard merchant ships. Occup. Environ. Med. 59, 85. https://doi.org/10.1136/oem.59.2.85.

Heath, C., Luff, P., 1991. Collaborative Activity and Technological Design: Task Coordination in London Underground Control Rooms. In: Bannon, L., Robinson, M.,

Schmidt, K. (Eds.), Proceedings of the Second European Conference on Computer-Supported Cooperative Work ECSCW '91. Springer Netherlands, Dordrecht, pp. 65–80. https://doi.org/10.1007/978-94-011-3506-1_5.
Hetherington, C., Flin, R., Mearns, K., 2006. Safety in shipping: The human element. J. Saf. Res. 37, 401–411. https://doi.org/10.1016/j.jsr.2006.04.007.
Huang, Y., Chen, L., Negenborn, R.R., van Gelder, P.H.A.J.M., 2020. A ship collision avoidance system for human-machine cooperation during collision avoidance. Ocean Eng. 217, 107913 https://doi.org/10.1016/j.oceaneng.2020.107913.
Hutchins, E., 1995. Cognition in the Wild. MIT press.
IALA, 2019. Position on the development of aids to navigation services (No. C70-8.2.1.2). International Association of Marine Aids to Navigation and Lighthouse Authorities.
Im, N., Seo, J.-H., 2010. Ship manoeuvring performance experiments using a free running model ship. TransNav Int. J. Marine Navigation Saf. od Sea Transp. 4.
IMO, 2021. Outcome of the Regulatory Scoping Exercise for the Use of Maritime Autonomous Surface Ships (MASS) (No. MSC.1/Circ.1638). London, UK.
IMO, 2019. Maritime Safety [WWW Document]. URL https://www.imo.org/en/OurWork/Safety/Pages/default.aspx (accessed 4.27.21).
IMO, 2018. IMO takes first steps to address autonomous ships [WWW Document]. URL http://www.imo.org/en/MediaCentre/PressBriefings/Pages/08-MSC-99-MASS-scoping.aspx (accessed 9.9.20).
IMO, 2017. International Convention on Standards ofTraining, Certification and Watchkeeping for Seafarers (STCW) 1978, as amended in 1995/2010.
IMO, 2013. Guidelines for the approval of alternatives and equivalents as provided for in various IMO instruments (No. MSC.1/Circ.1455). IMO, London, UK.
Kari, R., Steinert, M., 2021. Human factor issues in remote ship operations: Lesson learned by studying different domains. J. Marine Sci. Eng. 9 https://doi.org/10.3390/jmse9040385.
Katsivela, M., 2020. Unmanned Vessels and Regulatory Concerns. J. Int. Maritime Law.
Kim, T., Mallam, S., 2020. A Delphi-AHP study on STCW leadership competence in the age of autonomous maritime operations. WMU J. Maritime Affairs 19, 163–181.
Kimball, P., Bailey, J., Das, S., Geyer, R., Harrison, T., Kunz, C., Manganini, K., Mankoff, K., Samuelson, K., Sayre-McCord, T., Straneo, F., Traykovski, P., Singh, H., 2014. The WHOI Jetyak: An autonomous surface vehicle for oceanographic research in shallow or dangerous waters. In: 2014 IEEE/OES Autonomous Underwater Vehicles (AUV). pp. 1–7. https://doi.org/10.1109/AUV.2014.7054430.
Klein, G.A., 2017. Sources of power: How people make decisions. MIT press.
Kobayashi, H., Kohshima, S., 2001. Unique morphology of the human eye and its adaptive meaning: comparative studies on external morphology of the primate eye. J. Hum. Evol. 40, 419–435. https://doi.org/10.1006/jhev.2001.0468.
Kobayashi, H., Kohshima, S., 1997. Unique morphology of the human eye. Nature 387, 767–768. https://doi.org/10.1038/42842.
Kongsberg, 2020a. Autonomous ship project, key facts about YARA Birkeland [WWW Document]. URL https://www.kongsberg.com/maritime/support/themes/autonomous-ship-project-key-facts-about-yara-birkeland/ (accessed 11.8.21).
Kongsberg, 2020b. First adaptive transit on Bastøfosen VI [WWW Document]. URL https://www.kongsberg.com/maritime/about-us/news-and-media/news-archive/2020/first-adaptive-transit-on-bastofosen-vi/ (accessed 4.27.21).
Kooij, C., Hekkenberg, R., 2021. Identification of a task-based implementation path for unmanned autonomous ships. Maritime Policy Manage. https://doi.org/10.1080/03088839.2021.1914878.
Låg, S., 2019. ROMAS - Remote Operations of Machinery and Automation Systems - experiences from pilot testing.
Lieder, F., Griffiths, T.L., 2020. Resource-rational analysis: Understanding human cognition as the optimal use of limited computational resources. Behav. Brain Sci. 43, e1 https://doi.org/10.1017/S0140525X1900061X.
Lloyd's Register, 2016. Cyber-enabled ships: ShipRight procedure - autonomous ships (Guidance Document). Lloyd's Register Group, Southampton, UK.
Lunde-Hanssen, L.S., Braseth, A.O., Strand, S., 2020. Identification of information requirements in ROC operations room (No. IFE/E-2020/007). IFE, Halden, Norway.
Lützhöft, M.H., Dekker, S.W.A., 2002. On Your Watch: Automation on the Bridge. J. Navig. 55, 83–96. https://doi.org/10.1017/S0373463301001588.
Man, Y., Weber, R., Cimbritz, J., Lundh, M., MacKinnon, S.N., 2018. Human factor issues during remote ship monitoring tasks: An ecological lesson for system design in a distributed context. Int. J. Ind. Ergon. 68, 231–244. https://doi.org/10.1016/j.ergon.2018.08.005.
McCarthy, J., Minsky, M.L., Rochester, N., Shannon, C.E., 2006. A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence, August 31, 1955. AIMag 27, 12. https://doi.org/10.1609/aimag.v27i4.1904.
MiT, 2020. Roboat project [WWW Document]. roboat.org. URL http://www.roboat.org (accessed 11.19.20).
Moher, D., Liberati, A., Tetzlaff, J., Altman, D.G., 2009. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS med. 6, e1000097.
Moravec, H., 1988. Mind children: The future of robot and human intelligence. Harvard University Press.
Munim, Z.H., Saha, R., Schøyen, H., Ng, A.K., Notteboom, T.E., 2021. Autonomous ships for container shipping in the Arctic routes. J. Mar. Sci. Technol. 1–15.
Akhtar, N., Mian, A., 2018. Threat of Adversarial Attacks on Deep Learning in Computer Vision: A Survey. IEEE Access 6, 14410–14430. https://doi.org/10.1109/ACCESS.2018.2807385.
National Transportation Safety Board, 2020. Collision Between a Sport Utility Vehicle Operating With Partial Driving Automation and a Crash Attenuator, Mountain View, California, March 23, 2018 (Accident Report No. NTSB/HAR-20/01). National Transportation Safety Board, Washington, D.C., USA.
Nicas, J., Kitroeff, N., Gelles, D., Glanz, J., 2019. Boeing built deadly assumptions into 737 Max, blind to a late design change. The New York Times.
NTNU, 2021. SFI AutoShip - NTNU [WWW Document]. URL https://www.ntnu.edu/sfi-autoship (accessed 3.21.21).
NVivo, 2020. . QSR International.
Nzengu, W., Faivre, J., Pauwelyn, A.-S., Bolbot, V., Lien Wennersberg, L.A., Theotokatos, G., 2021. Regulatory framework analysis for the unmanned inland waterway vessel. WMU J. Maritime Affairs. https://doi.org/10.1007/s13437-021-00237-z.
Paavola, S., 2006. On the origin of ideas: An abductivist approach to discovery.
Parasuraman, R., Manzey, D.H., 2010. Complacency and bias in human use of automation: An attentional integration. Hum. Factors 52, 381–410.
Parasuraman, R., Riley, V., 1997. Humans and Automation: Use, Misuse, Disuse, Abuse. Hum. Factors 39, 230–253. https://doi.org/10.1518/001872097778543886.
Peeters, G., Kotzé, M., Afzal, M.R., Catoor, T., Van Baelen, S., Geenen, P., Vanierschot, M., Boonen, R., Slaets, P., 2020a. An unmanned inland cargo vessel: Design, build, and experiments. Ocean Eng. 201, 107056.
Peeters, G., Yayla, G., Catoor, T., Van Baelen, S., Afzal, M.R., Christofakis, C., Storms, S., Boonen, R., Slaets, P., 2020b. An Inland Shore Control Centre for Monitoring or Controlling Unmanned Inland Cargo Vessels. J. Marine Sci. Eng. 8, 758.
Zaccone, R., Martelli, M., Figari, M., 2019. A COLREG-Compliant Ship Collision Avoidance Algorithm. In: European Control Conference (ECC). Presented at the 2019 18th European Control Conference (ECC), pp. 2530–2535.
Ramos, M., Thieme, C.A., Utne, I.B., Mosleh, A., 2020a. Human-system concurrent task analysis for maritime autonomous surface ship operation and safety. Reliab. Eng. Syst. Saf. 195, 106697.
Ramos, M., Utne, I.B., Mosleh, A., 2019. Collision avoidance on maritime autonomous surface ships: Operators' tasks and human failure events. Saf. Sci. 116, 33–44. https://doi.org/10.1016/j.ssci.2019.02.038.
Ramos, M.A., Thieme, C.A., Utne, I.B., Mosleh, A., 2020b. A generic approach to analysing failures in human–System interaction in autonomy. Saf. Sci. 129, 104808.
Rasmussen, J., 1997. Risk management in a dynamic society: a modelling problem. Saf. Sci. 27, 183–213. https://doi.org/10.1016/S0925-7535(97)00052-0.
Reddy, N.P., Zadeh, M.K., Thieme, C.A., Skjetne, R., Sorensen, A.J., Aanondsen, S.A., Breivik, M., Eide, E., 2019. Zero-Emission Autonomous Ferries for Urban Water Transport: Cheaper, Cleaner Alternative to Bridges and Manned Vessels. IEEE Electrif. Mag. 7, 32–45.
Relling, T., Lützhöft, M., Ostnes, R., Hildre, H.P., 2021. The contribution of Vessel Traffic Services to safe coexistence between automated and conventional vessels. Maritime Policy Manage. https://doi.org/10.1080/03088839.2021.1937739.
Reyhanoglu, M., 1997. Exponential stabilization of an underactuated autonomous surface vessel. Automatica 33, 2249–2254. https://doi.org/10.1016/S0005-1098(97)00141-6.
Rødseth, Ø.J., 2017. Definitions for Autonomous Merchant Ships 22.
Rødseth, Ø.J., Burmeister, H.-C., 2015. Risk assessment for an unmanned merchant ship. TransNav Int. J. Marine Navigation Safety Od Sea Transp. 9.
Rødseth, Ø.J., Kvamstad, B., Porathe, T., Burmeister, H.-C., 2013. Communication architecture for an unmanned merchant ship. In: Presented at the 2013 MTS/IEEE OCEANS-Bergen, pp. 1–9.
Rødseth, Ø.J., Wennersberg, L.A.L., Nordahl, H., 2021. Towards approval of autonomous ship systems by their operational envelope. J. Mar. Sci. Technol. 1–10.
Røed, W., Mosleh, A., Vinnem, J.E., Aven, T., 2009. On the use of the hybrid causal logic method in offshore risk analysis. Reliab. Eng. Syst. Saf. 94, 445–455. https://doi.org/10.1016/j.ress.2008.04.003.
Rolls-Royce, 2018. Press releases [WWW Document]. Rolls-Royce and Finferries demonstrate world's first Fully Autonomous Ferry. URL https://www.rolls-royce.com/media/press-releases.aspx (accessed 4.18.21).
Rothblum, A.M., 2000. Human error and marine safety. Presented at the National Safety Council Congress and Expo, U.S. Coast Guard Research & Development Center, Orlando, FL, USA.
Rutledal, D., Relling, T., Resnes, T., 2020. It's not all about the COLREGs: a case-based risk study for autonomous coastal ferries. In: Presented at the IOP Conference Series: Materials Science and Engineering, IOP Publishing, p. 012016.
SAE International, 2017. Taxonomy and Definitions for Terms Related to On-Road Motor Vehicle Automated Driving Systems. SAE International. https://doi.org/10.4271/J3016_201401.
Saha, R., 2021. Mapping competence requirements for future shore control center operators. Maritime Policy Manage. https://doi.org/10.1080/03088839.2021.1930224.
Sharma, A., Kim, T., 2021. Exploring technical and non-technical competencies of navigators for autonomous shipping. Maritime Policy Manage. https://doi.org/10.1080/03088839.2021.1914874.
Shneiderman, B., 2020. Human-Centered Artificial Intelligence: Reliable. Safe & Trustworthy. null 36, 495–504. https://doi.org/10.1080/10447318.2020.1741118.
Shneiderman, B., 2016. Opinion: The dangers of faulty, biased, or malicious algorithms requires independent oversight. Proc. Natl. Acad. Sci. USA 113, 13538–13540. https://doi.org/10.1073/pnas.1618211113.
Statheros, T., Howells, G., Maier, K.M., 2008. Autonomous Ship Collision Avoidance Navigation Concepts, Technologies and Techniques. J. Navig. 61, 129–142. https://doi.org/10.1017/S037346330700447X.
Stone, P., Brooks, R., Brynjolfsson, E., Calo, R., Etzioni, O., Hager, G., Hirschberg, J., Kalyanakrishnan, S., Kamar, E., Kraus, S., 2016. Artificial intelligence and life in 2030: the one hundred year study on artificial intelligence.
Størkersen, K.V., 2020. Safety management in remotely controlled vessel operations. Marine Policy. https://doi.org/10.1016/j.marpol.2020.104349.

Sullivan, B.P., Desai, S., Sole, J., Rossi, M., Ramundo, L., Terzi, S., 2020. Maritime 4.0 – Opportunities in Digitalization and Advanced Manufacturing for Vessel Development. Procedia Manuf. 42, 246–253. https://doi.org/10.1016/j.promfg.2020.02.078.

Swedberg, R., 2014. The art of social theory. Princeton University Press.

Sheridan, T.B., 2011. Adaptive Automation, Level of Automation, Allocation Authority, Supervisory Control, and Adaptive Control: Distinctions and Modes of Adaptation. IEEE Trans. Syst. Man Cybern. - Part A: Syst. Hum. 41, 662–667. https://doi.org/10.1109/TSMCA.2010.2093888.

Tam, C., Bucknall, R., Greig, A., 2009. Review of collision avoidance and path planning methods for ships in close range encounters. J. Navigation 62, 455.

Tesla, 2019. Tesla Vehicle Safety Report [WWW Document]. Tesla Vehicle Safety Report. URL https://www.tesla.com/VehicleSafetyReport (accessed 4.22.21).

Tetley, W., 2008. Marine Cargo Claims. Yvon Blais 877–888.

Thieme, C.A., Utne, I.B., 2017. Safety performance monitoring of autonomous marine systems. Reliab. Eng. Syst. Saf. 159, 264–275.

Thieme, C.A., Utne, I.B., Haugen, S., 2018. Assessing ship risk model applicability to Marine Autonomous Surface Ships. Ocean Eng. 165, 140–154.

Tomasello, M., Hare, B., Lehmann, H., Call, J., 2007. Reliance on head versus eyes in the gaze following of great apes and human infants: the cooperative eye hypothesis. J. Hum. Evol. 52, 314–320. https://doi.org/10.1016/j.jhevol.2006.10.001.

United Nations, 1982. United Nations Convention on the Law of the Sea (UNCLOS).

U.S. National Highway Traffic Safety Administration, 2015. Critical Reasons for Crashes Investigated in the National Motor Vehicle Crash Causation Survey (No. DOT HS 812 115). U.S. Department of Transportation, Washington, D.C., USA.

Utne, I.B., Rokseth, B., Sørensen, A.J., Vinnem, J.E., 2020. Towards supervisory risk control of autonomous ships. Reliab. Eng. Syst. Saf. 196, 106757 https://doi.org/10.1016/j.ress.2019.106757.

Utne, I.B., Sørensen, A.J., Schjølberg, I., 2017. Risk Management of Autonomous Marine Systems and Operations. In: Presented at the ASME 2017 36th International Conference on Ocean, Offshore and Arctic Engineering, American Society of Mechanical Engineers Digital Collection. https://doi.org/10.1115/OMAE2017-61645.

Vagia, M., Rødseth, Ø.J., 2019. A taxonomy for autonomous vehicles for different transportation modes. Presented at the Journal of Physics: Conference Series, IOP Publishing, p. 012022.

Vagia, M., Transeth, A.A., Fjerdingen, S.A., 2016. A literature review on the levels of automation during the years. What are the different taxonomies that have been proposed? Appl. Ergon. 53, 190–202. https://doi.org/10.1016/j.apergo.2015.09.013.

Valdez Banda, O.A., Kannos, S., Goerlandt, F., van Gelder, P.H.A.J.M., Bergström, M., Kujala, P., 2019. A systemic hazard analysis and management process for the concept design phase of an autonomous vessel. Reliab. Eng. Syst. Saf. 191 https://doi.org/10.1016/j.ress.2019.106584.

Varela, F.J., Thompson, E., Rosch, E., 2016. The embodied mind: Cognitive science and human experience. MIT press, Cambridge, MA.

Vasstein, K., Brekke, E.F., Mester, R., Eide, E., 2020. Autoferry Gemini: a real-time simulation platform for electromagnetic radiation sensors on autonomous ships. IOP Conf. Ser.: Mater. Sci. Eng. 929, 012032 https://doi.org/10.1088/1757-899x/929/1/012032.

Wang, W., Gheneti, B., Mateos, L.A., Duarte, F., Ratti, C., Rus, D., 2019. Roboat: An Autonomous Surface Vehicle for Urban Waterways. In: in: 2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp. 6340–6347. https://doi.org/10.1109/IROS40897.2019.8968131.

Waymo, 2020. Waymo Safety Report. Mountain View, CA, USA.

Weick, K.E., 1976. Educational organizations as loosely coupled systems. Adm. Sci. Q. 1–19.

Wróbel, K., 2021. Searching for the origins of the myth: 80% human error impact on maritime safety. Reliab. Eng. Syst. Saf. 107942.

Wróbel, K., Gil, M., Chae, C.-J., 2021. On the influence of human factors on safety of remotely-controlled merchant vessels. Appl. Sci. (Switzerland) 11, 1–17. https://doi.org/10.3390/app11031145.

Wróbel, K., Gil, M., Montewka, J., 2020. Identifying research directions of a remotely-controlled merchant ship by revisiting her system-theoretic safety control structure. Saf. Sci. 129, 104797.

Wróbel, K., Montewka, J., Kujala, P., 2018. System-theoretic approach to safety of remotely-controlled merchant vessel. Ocean Eng. 152, 334–345.

Wróbel, K., Montewka, J., Kujala, P., 2017. Towards the assessment of potential impact of unmanned vessels on maritime transportation safety. Reliab. Eng. Syst. Saf. 165, 155–169.

Wu, B., Li, G., Wang, T., Hildre, H.P., Zhang, H., 2021a. Sailing status recognition to enhance safety awareness and path routing for a commuter ferry. Ships Offshore Struct. 1–12.

Wu, B., Li, G., Zhao, L., Aandahl, H.-I.J., Hildre, H.P., Zhang, H., 2021b. Navigating patterns analysis for on-board guidance support in crossing collision avoidance operations.

Wu, Q., Wang, T., Diaconeasa, M.A., Mosleh, A., Wang, Y., 2020. A Comparative Assessment of Collision Risk of Manned and Unmanned Vessels. J. Marine Sci. Eng. 8, 852.

Yara, 2017. YARA and KONGSBERG enter into partnership to build world's first autonomous and zero emissions ship | Yara International [WWW Document]. URL https://www.yara.com/corporate-releases/yara-and-kongsberg-enter-into-partnership-to-build-worlds-first-autonomous-and-zero-emissions-ship/ (accessed 11.8.21).

Yoo, Y., Lee, J.-S., 2021. Collision Risk Assessment Support System for MASS RO and VTSO Support in Multi-Ship Environment of Vessel Traffic Service Area. J. Marine Sci. Eng. 9, 1143.

Yoshida, M., Shimizu, E., Sugomori, M., Umeda, A., 2021. Identification of the relationship between maritime autonomous surface ships and the operator's mental workload. Appl. Sci. (Switzerland) 11, 1–23. https://doi.org/10.3390/app11052331.

Yoshida, M., Shimizu, E., Sugomori, M., Umeda, A., 2020. Regulatory Requirements on the Competence of Remote Operator in Maritime Autonomous Surface Ship: Situation Awareness, Ship Sense and Goal-Based Gap Analysis. Appl. Sci. 10, 8751.

Zhang, M., Zhang, D., Yao, H., Zhang, K., 2020. A probabilistic model of human error assessment for autonomous cargo ships focusing on human–autonomy collaboration. Saf. Sci. 130, 104838.

Zhou, X.-Y., Liu, Z.-J., Wang, F.-W., Wu, Z.-L., 2021. A system-theoretic approach to safety and security co-analysis of autonomous ships. Ocean Eng. 222, 108569 https://doi.org/10.1016/j.oceaneng.2021.108569.