Experimental trust dynamics modelling in supervised autonomous ship navigation in collision avoidance scenarios

Song, Rongxin; Papadimitriou, Eleonora; Negenborn, Rudy R.; van Gelder, Pieter

DOI
10.1016/j.trip.2025.101634

Publication date
2025

Document Version
Final published version

Published in
Transportation Research Interdisciplinary Perspectives

Transportation Research Interdisciplinary Perspectives 34 (2025) 101634

Contents lists available at ScienceDirect

Transportation Research Interdisciplinary Perspectives

journal homepage: www.sciencedirect.com/journal/transportation-research-interdisciplinary-perspectives

Experimental trust dynamics modelling in supervised autonomous ship navigation in collision avoidance scenarios

Rongxin Song [a,b,*], Eleonora Papadimitriou [b], Rudy R. Negenborn [c], Pieter van Gelder [b]

[a] Intelligent Transportation Systems Research Center, Wuhan University of Technology, China
[b] Safety and Security Science Group, Delft University of Technology, the Netherlands
[c] Department of Maritime and Transport Technology, Delft University of Technology, the Netherlands

ARTICLE INFO

Keywords:
Trust dynamics
Trust model
Trust Bayesian Network (TBN)
MASS under supervision
Mixed waterborne transport system (MWTS)
Linear mixed model (LMM)
Collision avoidance

ABSTRACT

Maritime Autonomous Surface Ships (MASS) are advancing the shipping industry, requiring a mixed waterborne transport system (MWTS) where human supervision provides a supporting role for maintaining safety and efficiency, particularly in complex scenarios. This study explores the dynamics of seafarers' trust in MASS during collision avoidance (CA) scenarios involving a vessel approaching from the starboard side. An empirical study with 26 participants representing diverse maritime experience levels examined how time, demographic factors, and collision avoidance strategies influence trust. Using a linear mixed model (LMM), trust was found to fluctuate across navigation stages: gradual accumulation during the routine navigation stage, sharp dissipation during strategy determination and execution stages, and partial recovery at the final stage. Strategies aligned with maritime regulations and appropriately timed evasive actions fostered higher trust, while overly early or imminent actions reduced trust. Additionally, a factor analysis consolidated the five trust dimensions, including dependability, predictability, anthropomorphism, faith, and safety, into two aspects: System Competence, encompassing the first four dimensions, and Situational Safety, representing safety-related trust. Furthermore, Bayesian Network (BN) is developed to model trust in the autonomous decision-making of MASS, integrating human observers demographics and situational factors. The model captures sequential trust dependencies, revealing the cascading effects of trust across various stages and the role of System Competence in shaping overall trust in the entire decision-making process. These findings provide actionable insights for designing MASS that support trust-building and optimise collision avoidance strategies, contributing to safer and more efficient autonomous maritime operations.

Introduction

Background

Maritime Autonomous Surface Ships (MASS) are being increasingly recognised for their potential to enhance operational efficiency and safety within the maritime industry. While advances in automation technology lay the groundwork, it is primarily the integration of intelligent systems that enable ships to perform navigation tasks autonomously, reducing the need for constant human control. However, human supervision will remain important in the near future (Negenborn et al., 2023), as autonomous systems may require monitoring and necessary intervention to ensure safe operations, especially in complicated navigational environments, e.g., for collision avoidance (CA) scenarios. In such situations, human supervisors are important in overseeing the system's actions and intervening when necessary.

Trust in autonomous systems is a key factor in ensuring safe and efficient collaboration between autonomous systems of MASS and human operators (Song et al., 2024b). A proper level of trust facilitates human operators to confidently delegate navigational tasks to these systems in specific scenarios. Trust affects how operators perceive the system's actions, their willingness to rely on the system, and their readiness to intervene when required. In addition, trust is not static (Kirkpatrick et al., 2017); it fluctuates based on factors such as system performance, environmental conditions, and operator characteristics (Poornikoo et al., 2024). Understanding the dynamics of trust in MASS operations, particularly with CA scenarios, is foundational for developing systems that maintain suitable trust levels.

* Corresponding author at: Safety and Security Science Group, Delft University of Technology, Delft, the Netherlands.
E-mail addresses: r.song@whut.edu.cn, r.song@tudelft.nl (R. Song).

https://doi.org/10.1016/j.trip.2025.101634
Received 27 January 2025; Received in revised form 14 July 2025; Accepted 12 September 2025
Available online 18 September 2025

Despite advances in automation, the dynamics of trust in human-supervised autonomous ship navigation, especially in CA contexts, remain underexplored. In CA scenarios, compliance with the Convention on the International Regulations for Preventing Collisions at Sea (COLREGs) is legally required for safe navigation. Operator trust in autonomous systems likely depends on how consistently the system adheres to these regulations, particularly in terms of its evasion strategies and the timing of evasive actions. Studying these dynamics, however, is challenging due to the limited availability of real-world interaction data between MASS and manned vessels, which restricts empirical insights and limits data-driven model development.

Recent studies have highlighted that one major challenge in MASS collision avoidance (CA) lies in the vagueness and interpretative flexibility of the COLREGs. (Wróbel et al., 2022) discussed that the current state of collision avoidance systems struggles with fully adhering to COLREG requirements due to rule vagueness. (Chang et al., 2024) conducted a systematic review, identifying gaps in aligning MASS decision-making with COLREG frameworks across varied navigational contexts. Additionally, concepts such as "declarative ship arenas" (Zarzycki et al., 2025) and "critical danger areas" (Gil, 2021) have been proposed to provide clearer spatial and temporal guidance for MASS avoidance actions. These works underline the necessity of exploring how such operational uncertainties influence human trust in MASS decision-making under CA scenarios.

Given these constraints, the central research problem addressed by this study is: *How can operator trust in MASS be measured, analysed, and modelled within controlled experimental settings*? Addressing this problem is important for understanding how operators interact with MASS, as trust dynamics directly influence operator confidence, intervention likelihood, and overall system performance.

In this study, the supervised autonomous vessel refers to systems where human operators are responsible for both monitoring and potential intervention. This study specifically focuses on the observational aspect of human supervision. Participants acted as observers, assessing the autonomous system's collision avoidance decisions without engaging in direct control actions.

To address this problem, this study makes the following contributions:

1. Design simulator-based experiments to simulate CA scenarios between MASS and conventional ships, enabling the controlled collection of trust-related data.
2. Explore how operator trust varies over time in CA situations through a linear mixed model (LMM), identifying and quantifying the influence of key factors, such as evasion strategies and timings, on trust dynamics.
3. Develop a Trust Bayesian Network (TBN) model to further analyse and predict operator trust dynamics in CA scenarios, focusing on diagnostic analysis informed by sensitivity analysis and predictive reasoning.

Structure of the paper

This paper is structured as follows: Section 2 reviews existing theories for trust investigation and modelling in human-robot/autonomous vehicles interaction, explicitly focusing on maritime applications. Section 3 details the experiment scheme and the framework for analysing human trust dynamics. Section 4 presents the development of the TBN model, describing how the model is constructed to predict human trust based on empirical data and key influencing factors. Section 5 discusses the results of the empirical study and the evaluation of the TBN model. Section 6 concludes the research and provides directions for future research.

State of the art

In recent years, the study of human trust in human-autonomy interaction has gained much attention, particularly in critical domains such as autonomous navigation (Basu and Singhal, 2016). The reason is that trust influences the safety and efficiency of these interactions through its effect on operator behaviour: appropriate trust reduces unnecessary intervention while maintaining adequate oversight. In this section, we will explore the nature of trust in the human-autonomy interaction, the methods used to measure and investigate trust, and various approaches to modelling trust.

Trust in human-autonomy interaction

Nature of trust

In human-autonomy interaction, trust is commonly defined as a user's willingness to be vulnerable to the actions of an autonomous system based on positive expectations of its performance. A widely referenced conceptualisation of trust was proposed in (Mayer et al., 1995), which characterises trustworthiness through three critical dimensions: ability, benevolence, and integrity. In this model, ability refers to the system's competence in fulfilling tasks, benevolence captures the alignment of the system's goals with those of the user, and integrity reflects the system's adherence to acceptable standards.

In (Lee and See, 2004), a definition of trust drawn from previous studies is the attitude that an agent will help achieve an individual's goals in a situation characterised by uncertainty and vulnerability. It is pointed out that proper trust calibration prevents overtrust (misuse) and undertrust (disuse) by ensuring that user trust corresponds to the system's real-world performance. Furthermore, this definition was used by (Guo and Yang, 2021) to investigate the evolution of trust within human–computer interaction, categorising users into Bayesian decision-makers, oscillators, and disbelievers, each reflecting unique patterns of trust adjustment.

Expanding further, focusing on the factors that may have an impact on trust, (Peter A. Hancock et al. 2011) presented a *meta*-analytic framework that provides an empirical perspective by examining human, robot, and environmental factors affecting trust in human-robot interaction. Their *meta*-analysis concludes that robot performance and attribute-based factors are significant contributors to trust development, while environmental factors play a moderate role.

Trust is dynamic and responsive to changes in the operational environment (PARK et al., 2008) and system performance (Alhaji et al., 2023). Moreover, dynamic models, such as OPTIMo proposed by (Xu and Dudek 2015a), conceptualise trust as a probabilistic and context-sensitive belief that adapts in real-time to fluctuations in system performance. Trust is viewed as continually updated based on system behaviour, contrasting with static measures that provide only a momentary view.

A three-layered trust model comprising dispositional, situational, and learned trust was proposed by (Hoff and Bashir, 2015) to better understand trust in human-automation interactions. In this model, dispositional trust is an individual's inherent tendency to trust or distrust automation, situational trust arises from contextual elements like task complexity and perceived risk, and learned trust builds through prior experiences with the system. This layered approach integrates individual, contextual, and experiential factors, illustrating that trust varies independently across these layers and is influenced by the dynamic interplay between user expectations and real-time system feedback.

Recently, frameworks like IMPACTS proposed by (Hou et al., 2021) have extended these trust considerations to encompass practical characteristics essential for building trust in autonomous systems. The model identifies seven characteristics, including intention, measurability, performance, adaptivity, communication, transparency, and security, as crucial for establishing and sustaining trust in autonomy. This model emphasises adaptability and real-time feedback mechanisms, aligning with dynamic models while providing actionable insights for designing trust-supportive systems. Its practical relevance is notable in high-stakes domains, where decision-making must be precise, transparent, and adaptive to changing conditions, underscoring trust as a dynamic, context-sensitive construct.

Additionally, key factors that may influence trust evolvement were investigated by (Alhaji et al., 2021), including reliability, predictability, and dependability. Further, studies by (Alhaji et al., 2023) focused on the accumulation and decay of trust, identifying that trust can be asymmetrical in response to system performance: while reliability is crucial in building trust, its erosion is more pronounced when systems fail, particularly in high-risk environments.

In summary, while existing studies provide insights into the nature, dimensions, and dynamics of trust in human-autonomy interaction, their application to MASS remains limited, particularly during critical operational scenarios like collision avoidance. In this study, we address this gap by incorporating both the dynamic and its multidimensional characteristics, such as reliability, predictability, and safety. Using these established theories, we aim to understand trust dynamics and characteristics in MASS in CA scenarios.

Trust measurement and investigation

In the study of trust dynamics within human-autonomy interaction, researchers utilise a diverse array of measurement methods, including subjective, objective, and hybrid techniques.

Subjective methods, such as self-report questionnaires (Malle and Ullman, 2020), allow operators to directly express their perceived trust levels. For instance, frequent trust measurement intervals have been used to observe how trust levels shift in response to interaction quality and timing (Jackson et al., 2022). A subjective trust measurement scale tailored to human-robot interaction was developed and validated by (Yagoda and Gillan, 2012), exploring how dispositional and history-based trust components influence user trust in varying contexts.

In contrast, objective methods provide physiological indicators of trust fluctuations during task execution. For instance, using psycho-physiological data, such as heart rate variability, electrodermal activity, and Electroencephalography (EEG), offers insights into trust dynamics within virtual environments by identifying immediate physiological responses associated with trust levels (Chauhan et al., 2024). Among these, EEG signals capture the brain's immediate response under conditions of trust and thus have been used as a more objective physiological indicator (Wang et al., 2018; Xu et al., 2024). In addition, eye tracking, another measurement method, has been employed to infer trust levels. For example, it was combined with Bayesian models to be used to estimate the workload of operators in real time (Luo et al., 2024).

To provide a more comprehensive view of trust, hybrid measurements have emerged, integrating both subjective and objective data. The study (Krausman et al., 2022) proposed a toolkit for trust measurement in human-autonomy teams, combining self-report, behavioural indicators (e.g., reliance, compliance, eye-tracking), and physiological data (e.g., heart rate variability) to capture dynamic trust levels. Furthermore, the study (Hopko et al., 2023) examined how cognitive fatigue, robot reliability, and operator gender impact trust in collaborative robots, where both physiological (performance, heart rate variability) and subjective measures (surveys) were employed.

In exploring trust dynamics, statistical methods are commonly employed to analyse how trust varies under different conditions. Techniques like ANOVA and Signal Temporal Logic (STL) are utilised to assess environmental impacts on trust, examining factors such as alarm types or task conditions (Sheng et al., 2019). Multi-factor analysis, including t-tests and correlation, further reveals how interaction levels and workspace settings impact trust, supporting a nuanced understanding of trust fluctuations (Chauhan et al., 2024). In addition, Linear Mixed Models have been instrumental in capturing trust dynamics over time. For example, exploring specific EEG frequencies (Delta and Gamma) associated with trust fluctuations (Wang et al., 2018) and investigating the variables of time and frequency, showing the accumulation effect of the frequency of positive interactions on trust (Jackson et al., 2022).

Overall, trust measurement methods include subjective, objective, and hybrid approaches, each with advantages and limitations. Subjective methods are straightforward but are prone to bias. Objective methods provide real-time insights but require complex tools. Hybrid methods are comprehensive but costly. Among statistical approaches, LMM excels at capturing dynamics while accounting for individual differences, whereas traditional methods like ANOVA are limited in handling repeated measures and complex hierarchical data. This study uses subjective measurements to gather trust data and apply LMM to analyse its dynamic evolution in MASS collision avoidance scenarios.

Trust computational models

In the field of trust modelling for human-autonomy interaction, research has developed multiple approaches to address the dynamic nature of human trust in autonomous systems, each categorised by distinct modelling techniques. Probabilistic models are widely applied in trust modelling. Guo and Yang (2021b) employed Bayesian inference with a Beta distribution to capture trust adjustments following successful or unsuccessful robotic tasks, emphasising time dependency and the impact of negative experiences. Their findings categorised users into types (e.g., rational, oscillating, disbeliever), enabling real-time trust updates tailored to individual preferences. The OPTIMo model by (Xu and Dudek, 2015) combines dynamic Bayesian Networks (DBN) with feedback to estimate trust continuously in high-risk, multi-task settings. In multi-robot environments, (Zheng et al., 2023) used Bayesian optimisation and state-space equations for trust modelling, applying Markov Chain Monte Carlo and Bayesian Optimization Experimental Design to enhance task allocation. Additionally, (Fooladi Mahani et al., 2021) explored trust in multi-robot settings using a DBN-based model with Boltzmann machines, parameterised by an Expectation-Maximization (EM) algorithm, which aids operators in trust allocation across multiple autonomous agents. These probabilistic models provide high interpretability and adaptability, making them ideal for the real-time demands of human-autonomy interaction operations.

Time-series models further deepen trust modelling by analysing historical trust trends, enabling accurate predictions of future trust levels in sustained human-robot collaboration. (Guo and Yang, 2021a) leveraged time-series data to study trust's self-reinforcing effects and sensitivity to negative feedback. (Sadrfaridpour et al., 2016) combined time-series methods with neural networks to dynamically adjust robot speed in response to human feedback.

Decision-theoretic models apply structured frameworks, such as Markov Decision Processes (MDP) and Partially Observable MDPs (POMDP), to manage trust by integrating trust as a decision variable in task optimisation. (Wu et al., 2017) used an MDP-based trust model to optimise trust through dynamic task allocation, aligning with the safety–critical needs of MASS operations. (Chen et al., 2018; Chen et al., 2020) built on this by treating trust as a hidden variable within a POMDP, enabling trust inference and decision optimisation.

Machine learning and hybrid models offer enhanced predictive power and flexibility for managing complex, nonlinear trust dynamics. (Soh et al., 2020) combined Recurrent Neural Networks and Gaussian processes to capture trust shifts across tasks, providing adaptability for multi-task contexts. In the customer experience domain, (Roy et al., 2024) integrated Partial Least Squares Structural Equation Modeling with Artificial Neural Networks to analyse trust's nonlinear effects, highlighting trust's role in complex and interactive settings. (Lee et al., 2021) employed sparse Gaussian processes and deep neural networks to estimate uncertainty in trust, making their model suitable for decision-making in complex environments. Together, machine learning models address the need for precision and responsiveness in trust modelling, enabling autonomous systems to adjust to diverse operator requirements effectively.

Trust modelling methods, including probabilistic models, time-series analyses, and machine-learning techniques, offer different strengths for capturing trust dynamics. Among these, Bayesian networks excel in representing trust evolution and real-time updates. In this study, a BN-based approach is used to model trust dynamics in MASS collision avoidance scenarios, enabling the integration of trust changes with system performance across navigation stages.

Trust consideration in MASS's navigation

Following general theories of human-robot interaction, trust in MASS demands particular consideration of the multi-stakeholder context and the dynamic nature of maritime environments. In both Remote Control Centres (RCCs) for fully autonomous ships and in human-autonomy collaborative navigation scenarios, trust is essential for operators who must rely on indirect data transmission and operational feedback without direct physical control (Misas et al., 2022; Song et al., 2024b).

Recent studies in MASS collision-avoidance increasingly incorporate human supervisory perspectives. (Huang, 2019) proposed a human–machine cooperation system emphasising transparent decision-making and operator intervention support, which indirectly relates to trust by improving operator awareness and perceived control. (van de Merwe et al., 2024) highlighted that effective human supervision requires continuous, clear information, an element foundational to fostering trust through improved situational awareness. Furthermore, by translating COLREGs into a machine-executable fuzzy expert system, the study (Bakdi and Vanem, 2022) enhances the transparency and predictability of MASS collision avoidance behaviour, which is essential for fostering human trust in remote monitoring and human–machine collaborative maritime operations.

It is emphasised by (Misas et al., 2022) that in RCC settings, trust is closely linked to the reliability of data transmission and cybersecurity, both of which are critical for maintaining the situational awareness needed for safe supervision. Therefore, it is crucial to maintain network security and ensure the reliability of information transmission. Additionally, the study conducted by (Gregor et al., 2023) observed that high levels of VR immersion may introduce complexities, such as increased motion sickness and slower situational awareness response times, which, if left unchecked, could impact operator trust and decision-making.

Furthermore, a Schema World Action Research Method (SWARM) was employed by (Lynch et al., 2023) and (Lynch et al., 2024) to explore the decision-making process of MASS operators in a remote monitoring centre and to analyse the impact of trust on their operations in conjunction with the Trust Module, revealing that trust in high-automation settings relies heavily on precise feedback and transparent system behaviour.

By combining both quantitative and qualitative methods (questionnaires, interviews, and technician logs), a mixed-methods approach was used by (Alsos et al., 2024) to triangulate findings on public trust and system performance, showing that trust can fluctuate based on perceived system reliability and interaction context.

A decision-making framework designed for MASS was given by (Song et al., 2024b), where human trust was considered a key element that influences situational awareness and safe navigation of the decisions made by autonomous systems. The trust module of this framework can be modulated by taking human reactions as input during the interaction between human operators and MASS.

Trust within MASS is also recognised as extending beyond individual operators, encompassing collective trust across a mixed waterborne system in which autonomous and conventional ships co-exist (Mallam et al., 2020). In such an environment, trust among stakeholders becomes essential to facilitate safe and coordinated operations. Additionally, research conducted by (Mallam et al., 2020) examined the changing role of human operators in autonomous maritime systems, noting that trust is influenced by operators' understanding and control over system decisions. Trust is presented as vital for system predictability and reliability, especially as traditional seafaring skills become less relevant.

While progress has been made in exploring trust within MASS systems, critical gaps remain, particularly regarding trust dynamics in collision avoidance scenarios. To address these gaps, this study employs an empirical approach to investigate trust dynamics and develops a BN-based model to capture the evolution of trust across different navigation stages. By incorporating trust's multidimensional characteristics into the dynamic model, this research seeks to provide insights into the understanding of trust in MASS decision-making in CA scenarios.

Trust data collection and dynamics analysis

Definitions: Human trust in the context of MASS's autonomous navigation systems can be narrowly defined as the belief that humans hold to the autonomous system's capability of situational awareness and appropriate task implementation (Song et al., 2024b). The trust is dynamic, evolving across different stages of navigation and influenced by factors such as compliance with COLREGs, decision-making strategies, and the timing of evasive actions. Furthermore, trust levels and evasion timings are defined as follows:

(1) **Trust Levels**: human trust in MASS reflects their confidence in the system's autonomous performance. Higher trust means a stronger observer's confidence in MASS's abilities to perform tasks successfully, while lower trust refers to more frequent manual checks and doubts about the capabilities of MASS's decision-making system.
(2) **Evasion timing**: it refers to the latency between the identification of a potential collision object and the initiation of an evasive manoeuvre by MASS. It is categorised into three key timing windows in this study, listed below:
   - **Standard**: a range where manoeuvres are typically expected to take place based on conventional practices and safety standards. This is a dynamic window that adjusts based on the operational context, allowing for sufficient time to assess the situation and respond appropriately.
   - **Early**: initiating manoeuvres earlier than typically expected, providing additional safety margin. This timing anticipates potential risks and acts before the standard window. This aligns with the concept of "declarative ship arenas" discussed in maritime collision avoidance literature (Zarzycki et al., 2025).
   - **Imminent**: the very last feasible moment when collision avoidance must be executed. This timing is used as a last resort when all prior opportunities to mitigate the situation have passed. This corresponds to"critical areas" as defined in prior studies (Gil, 2021).

Building on the foundational concepts of trust levels and evasion timing, this study formulates three hypotheses to examine the dynamics of human trust in MASS during CA scenarios:

Hypothesis 1 (H1): Human trust in MASS will fluctuate, including trust accumulation and dissipation, depending on the system's compliance with COLREGs rules, in particular Rules 15, 16, and 17, and the timing of evasive manoeuvres such as early and imminent moments.

Hypothesis 2 (H2): Right-turn evasion strategies will lead to a higher trust of the participant than left-turn strategies in the scenario where a vessel approaches from the starboard side, assuming the importance of COLREGs compliance.

Hypothesis 3 (H3): Early evasion actions and imminent action in general risk situations for the COLREGs-aware MASS with a "give-way" role will lead to lower trust levels of human observers, assuming the importance of proper evasive timing.

Experimental design

This study investigates the dynamics of observer trust in MASS during collision avoidance scenarios. The experiment was conducted in two phases to examine both the evolution of trust and the impact of different factors influencing trust, such as compliance with COLREGs and timing of evasive actions. Participants observed simulated scenarios and evaluated their trust levels in a controlled environment where MASS executed various collision avoidance strategies in response to an approaching vessel from the starboard side.

Participants: The experiment engaged 26 participants recruited through maritime channels, including captains and officers, ensuring diverse professional experience levels. Each participant voluntarily took part in the experiment, and all had prior experience with ship navigation but with various experiences ranging from < 5 years to > 8 years. The experiment took place over two phases, lasting approximately 30 min per participant. To ensure clarity in the experimental context, all participants were explicitly informed prior to the start of each scenario that the observed vessel was an MASS.

Apparatus: The experiment was conducted using the NT-PRO 5000 ship manoeuvring simulator, a full-task ship navigation simulator that provides a realistic maritime environment. The simulator setup includes radar, ECDIS, and ARPA systems consistent with equipment found on operational bridges, featuring high-fidelity graphical environments and hydrodynamic modelling that accurately replicate real-world ship behaviours under varying navigational and environmental conditions. The simulator was designed to replicate a standard open-sea navigation scenario where an autonomous vessel encounters a conventional vessel from the starboard side. Table 1 presents the initial parameters of both the own and surrounding ships. The ships were set to start each scenario with identical positions, speeds, and headings, with the surrounding ship serving as a constant movement while the own ship executing predefined CA strategies.

To simulate the decision-making capabilities of MASS, an experienced ship operator controlled the vessel behind the scenes, following predetermined decision-making logic that emulated the behaviour of an autonomous system. As shown in Fig. 1, the logic included two key strategies: (1) *Left-turn and right-turn manoeuvres performed at standard timing*, representing compliance and deviation from COLREGs. (2) *In the right-turn condition, additional strategies involving early and imminent evasive actions were implemented* to examine the effect of manoeuvre timing.

Fig. 1 shows the vessel trajectories for all experimental conditions. It demonstrates how the own ship responded to the surrounding ship's movements. The trajectories displayed in Fig. 1 represent only the active collision avoidance phase, concluding once the own ship has safely avoided the target ship. This reflects the experimental design, where scenarios ended after the avoidance action was completed, and no collisions occurred throughout the study. In Fig. 1, 'OS' (Own Ship) represents the MASS executing collision avoidance manoeuvres, and 'TS' (Target Ship) refers to the conventional vessel maintaining its course and speed according to COLREGs. The depicted left-turn and right-turn strategies by OS were designed to investigate trust dynamics under both COLREG-compliant and non-compliant scenarios. The traffic separation scheme is shown for context only and does not influence the experimental logic. These behaviours were based on the initial parameters outlined in Table 1.

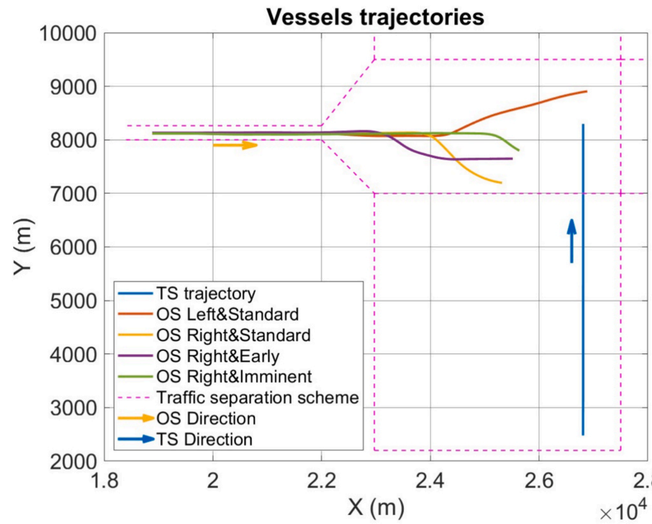


Fig. 1. Trajectories of the own and surrounding ships under varying conditions.

Participants' task in the experiment was to **observe** the scenarios and **evaluate** their trust levels in the MASS at various navigation and collision avoidance stages. They were instructed to focus on the system's decision-making behaviour, including evasive actions and timing. This approach ensured consistent and reproducible implementation of CA strategies. These included both COLREGs-compliant and non-compliant manoeuvres. The target vessel maintained a constant course and speed, in accordance with maritime regulations.

Experimental Design and Conditions: The experiment followed a two-phase structure, as shown in Fig. 2. The condition setting is presented in Table 2. The experiment consisted of two distinct phases: **Phase 1**, which explored trust levels associated with left-turn and right-turn strategies, and **Phase 2**, which examined trust differences between early and imminent manoeuvre responses in the right-turn condition. Participants were divided into two groups within each phase, experiencing the scenarios in reverse order.

Phase 1 –COLREGs compliance consideration: Participants observed the MASS navigating under two conditions: one in which the autonomous vessel complied with COLREGs by altering course to the starboard side to avoid the collision, and another where it neglected COLREGs with a left-turn strategy but still successfully avoided a collision. In both scenarios, MASS takes CA manoeuvres at standard timings, as previously defined, where manoeuvres are typically expected to take place based on conventional practices and safety standards. This design choice was made to control for timing-related variability when comparing compliant versus non-compliant strategies. Timing variations (early and imminent actions) were specifically examined in Phase 2 under the compliant right-turn condition.

Phase 2 – Evasion timing consideration: In the second phase, the focus was on the timing of CA strategies. Participants were exposed to two conditions: one where the MASS took early evasive action and another where it took imminent strategy.

In the two phases, trust dynamics were captured through post-scenario questionnaires after each run and were evaluated across five key stages, details presented below:

- ***Initial Trust***: At the beginning of the navigation process.
- ***Trust During Regular Navigation***: Before any collision-avoidance decisions are made.
- ***Trust During Decision-Making for Collision Avoidance***: As the ship initiates avoidance strategies.

Table 1
Initial parameters of the own and surrounding ships in the experimental scenarios.

Vessel	Ship type	Ship length	Width	Speed	Heading
The own vessel	Bulk carrier	225 *m*	32.3 *m*	10.5kn	090 °
The surrounding vessel	Container ship	190 *m*	30.0 *m*	12.0 kn	000 °

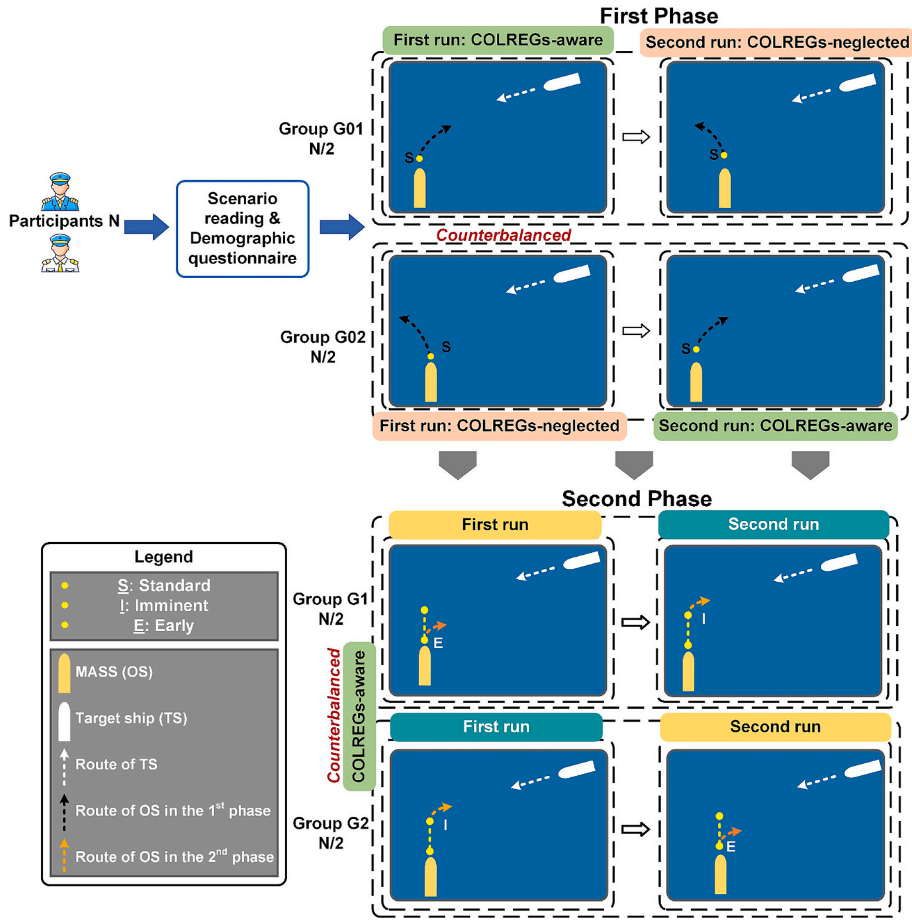


Fig. 2. Illustration of experimental procedure for collecting observers' trust in CA scenarios.

Table 2
Experimental groups and conditions. COLREGs-aware: succeeds with collision avoidance while complying with COLREGs. COLREGs-neglected: succeeds with collision avoidance but neglects COLREGs. Early: taking actions earlier than at the standard time that the corresponding action occurs. Imminent: taking imminent actions.

First phase – COLREGs compliance				Second phase – Timings		
Group No.	Trust dynamics			Group No.	Trust dynamics	
	First run	Second run			First run	Second run
G01 (N = 20)	COLREGs-aware	COLREGs-neglected	Break & randomisation	Group G1 (N = 20)	Early	Imminent
G02 (N = 20)	COLREGs-neglected	COLREGs-aware		Group G2 (N = 20)	Imminent	Early

- ***Trust During Collision-Avoidance Execution***: When the ship performs the manoeuvre after deciding on the CA strategy.
- ***Final Trust***: At the conclusion of the scenario, after the whole CA process has been completed.

Hereafter, these five stages are denoted as **Trust1** (Initial Trust), **Trust2** (Trust During Regular Navigation), **Trust3** (Trust During Collision-Avoidance Execution), **Trust4** (Trust During Collision-Avoidance Execution), and **Trust5** (Final Trust) for brevity and consistency in the subsequent analysis.

In addition to the stage-based trust assessments, trust was also measured across five key dimensions after each scenario, using specific questions designed to capture different aspects of trust. These dimensions were as follows:

- ***Dependability***: Assessed by asking participants to rate how confident they were in the MASS's ability to avoid collisions (e.g., "To what extent can you count on the MASS to avoid collisions in this scenario?").
- ***Predictability***: Evaluated based on how predictable the autonomous vessel's behaviour was according to standard maritime practices (e.g., "To what extent did you think the behaviour of the MASS was predictable based on standard maritime practices?").
- ***Anthropomorphism***: Related to the interpretation of non-human things or events in terms of human characteristics and measured by comparing the MASS's behaviour to that of a well-trained human operator (e.g., "How consistent was the MASS's behaviour with how a well-trained human operator would have acted?").
- ***Faith***: Captured by asking participants about their belief in the MASS's ability to handle future collision scenarios (e.g., "To what extent do you believe the MASS will be able to cope with all collision situations?").
- ***Safety***: Rated by asking how safe participants felt during the collision avoidance process (e.g., "How much do you feel unsafe in the whole process of autonomous collision avoidance?").

The trust questionnaire was adapted from validated scales in human-automation interaction (Yagoda and Gillan, 2012) and (Alhaji et al., 2021). It included five single-item 7-point Likert scales covering Dependability, Predictability, Anthropomorphism, Faith, and Safety. To ensure relevance in the MASS context, items were reviewed by maritime experts for content validity.

Data collection: Trust scores were collected through quantitative trust ratings in the post-scenario questionnaires administered via the *Qualtrics* platform, which allowed participants to reflect on their trust levels across various stages after each scenario. Trust scores across both the dynamic stages and dimensions were gathered, enabling further analysis of how trust evolved under different experimental conditions.

Procedure: participants were briefed on the experimental setup and provided with a demonstration of the ship manoeuvring simulator. A pre-experiment survey was administered to collect demographic information. After familiarising themselves with the simulator, participants proceeded with the scenarios in both phases. In each scenario, the participant observed an autonomous ship's behaviour varied according to the experimental conditions as the autonomous vessel encountered an approaching conventional vessel from the starboard side. The participant's view on engaging in simulator experiments is shown in Fig. 3.

After collecting the data, we conducted statistical analysis to compare trust dynamics. The analysis examined how trust changed over time and how it was influenced by different collision avoidance strategies and their timing.

Exploratory analysis

Data were collected via *Qualtrics* from a sample of 26 seafarers (hereafter referred to as "observers") with diverse backgrounds in terms of navigation experience, vessel types, positions, and age groups, enabling exploratory analysis of trust dynamics in autonomous navigation. The participants ranged from 29 to 55, with a majority falling between 30 and 35. In terms of position, the sample included captains (15.4 %), first officers (30.8 %), second officers (30.8 %), third officers (7.7 %), and pilots (15.4 %). Experience levels varied, with 50 % of participants reporting over eight years of maritime experience, 30.8 % between five to eight years, and 19.2 % with less than five years of experience. The types of vessels that the observer was familiar with were also diverse, including general vessels (65.4 %), tankers (23.1 %), and special-purpose vessels (11.5 %). Trust ratings were measured across five stages, with mean scores ranging from 3.62 to 3.94 (standard deviations of approximately 1.4 to 1.6).

Furthermore, a repeated measures analysis was conducted to investigate the dynamics of trust ratings across experimental stages. The result revealed significant variability in trust levels between participants, as indicated by the significant main effect of individual differences ($p < 0.001$). This result underscores the presence of underlying factors contributing to differences in trust across individuals. To further investigate the trust dynamics and account for both fixed effects (e.g., experimental conditions) and random effects (e.g., variability across participants), LMM was employed. This method is suitable for analysing repeated measures data while capturing individual differences.

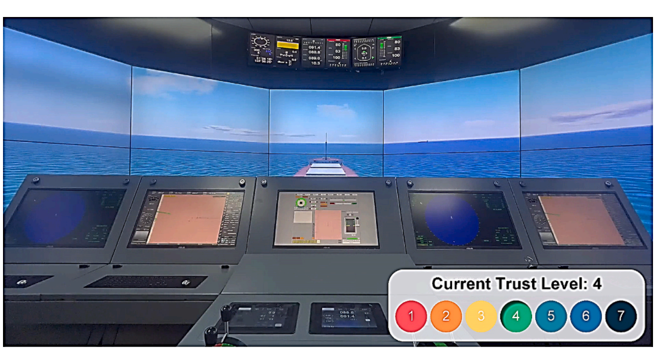


Fig. 3. Participants' view on engaging in simulator experiments.

LMM model development

Mann-Whitney U tests were first employed to evaluate trust differences across experimental orders within each phase to determine whether the order influenced participants' trust. The results indicated that for both the left–right strategy comparison and the early-imminent timing comparison, there were no significant differences in trust levels across any of the five measured trust dimensions (trust1 through trust5). Specifically, the p-values were 0.604 (Trust1), 0.672 (Trust2), 0.765 (Trust3), 0.443 (Trust4), and 0.852 (Trust5), all above the 0.05 threshold, suggesting that the sequence of presentation had no significant impact on trust ratings. Thus, the sequence of scenario presentation was considered to have a negligible impact on trust ratings. Consequently, sequence effects were excluded from the LMM to concentrate on primary factors of interest.

Given these findings, the LMM model includes the condition, trust moment (defined by the five key stages in Sec. 3.1), and demographic variables (e.g., experience, vessel type, position, age) as fixed effects, while individual participant differences were treated as random effects to account for variability in trust responses. Model performance was evaluated using multiple metrics. The model's marginal R^2 of 0.338 indicated that fixed effects alone explained 33.8 % of the variance, while the conditional R^2 reached 0.771, signifying that the combined influence of fixed and random effects accounted for 77.1 % of the overall variance. Subsequent analyses focus on significant main and interaction effects, providing insights into trust dynamics across various factors.

The statistical results of the main and interaction effects for trust are presented in Table 3, highlighting key factors influencing trust dynamics. The analysis reveals that trust moment and condition are significant predictors of trust, indicating that both the stages of navigation and the conditions influenced participants' trust in the autonomous system. Additionally, interaction effects between age, vessel type, and experience with trust moment suggest that trust evolved differently based on participants' maritime backgrounds and professional experience. These findings underscore the importance of operational context and individual characteristics in shaping trust, setting the stage for a more detailed examination of how these factors influence trust in autonomous navigation.

Main effects analysis

Fig. 4 presents the mean trust scores across five distinct stages of the navigation process, illustrating how trust levels evolve as the MASS progresses through various CA stages. Stage 1 (Initial Trust): Trust is measured at the outset, representing baseline confidence in the system before any navigation manoeuvres. Participants' trust at this stage serves as a reference level and shows relatively high stability. Stage 2 (Trust During Regular Navigation): Trust is assessed during standard navigation, prior to any collision-avoidance decisions. Here, trust levels remain close, with a slight increase to the initial levels, indicating that participants maintain a relatively steady trust during routine navigation without imminent risks. Stage 3 (Trust During Decision-Making for

Table 3
Type III Tests of Fixed Effects Dependent Variable: Trust.

Factors	Source	F	Sig.
Main effects	Intercept	137.462	<0.001
	Position	0.421	0.826
	Experience	1.558	0.245
	Vessel type	2.736	0.099
	Trust moment	2.840	**0.024**
	Condition	5.117	**0.002**
	Age	0.654	0.535
Interaction effects	Trust moment * age	5.723	**<0.001**
	Trust moment * condition	0.728	0.725
	Vessel type * trust moment	2.075	**0.037**
	Experience * trust moment	2.102	**0.034**
	Position * trust moment	0.653	0.872

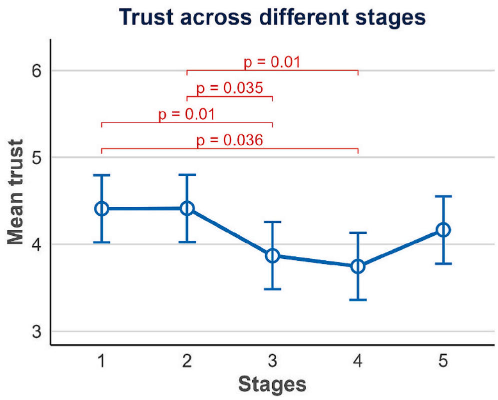


Fig. 4. Illustration of trust scores across all stages based on linear mixed models.

Collision Avoidance): Trust is recorded as the autonomous system initiates collision-avoidance strategies and timings. This stage shows a shape decline in trust compared to both Stage 1 (p = 0.01) and Stage 2 (p = 0.035), suggesting that participants' confidence diminishes when the system shifts from routine navigation to making critical decisions. Stage 4 (Trust During Collision-Avoidance Execution): Trust is further evaluated as the system performs the avoidance strategies. Another decline in trust is observed, with significant differences between Stage 1 and Stage 4 (p = 0.036) and Stage 2 and Stage 4 (p = 0.01), indicating increased participant uncertainty or caution during the strategy execution. Finally, at Stage 5 (Final Trust), Trust is assessed at the conclusion of the scenario after all manoeuvres have been executed. Trust levels partially recover at this stage but do not fully return to initial levels, suggesting residual caution even after observing the system's successful task completion.

Fig. 5 displays the mean trust levels across four conditions: Early/Starboard, Imminent/Starboard, Standard/Port, and Standard/Starboard. This comparison highlights how variations in collision-avoidance timing (early vs. imminent) and direction (starboard vs. port) affect trust in the autonomous system. A statistically significant difference between conditions is noted, with Standard/Starboard showing a higher mean trust than Standard/Port ($p < 0.001$).

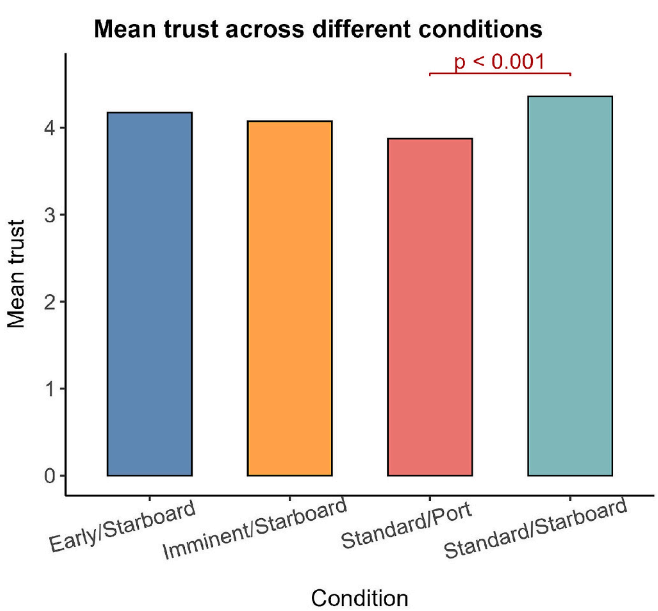


Fig. 5. Illustration of trust scores comparison between different conditions based on LMM.

Interaction effects analysis

As presented in Table 3, the significance test results indicate that trust dynamics vary significantly across navigation stages depending on observers' Vessel Type ($p = 0.037$), Experience Level ($p = 0.034$), and Age ($p < 0.001$). Given the lack of significance for other interactions, such as trust moment with *condition* ($p = 0.725$) and position ($p = 0.872$), the subsequent analysis focuses on these significant effects to provide a targeted exploration of trust dynamics across various stages. Thus, we analysed how trust scores varied across the five navigation stages (from initial to final trust) under specific demographic factors that have significant impacts. Fig. 6 illustrates these variations concerning three demographic variables: Vessel Type, Experience Level, and Age. Each subplot provides a focused view of how these demographic factors interact with trust dynamics, revealing distinct trends and potential influences at each stage.

For vessel type, as shown in Fig. 6(a), participants navigating tankers generally exhibited higher trust levels across all stages, while those associated with special-purpose ships showed a notable decline in trust from Stages 2 to 4.

In terms of experience level, as shown in Fig. 6(b), participants with less than 5 years of experience displayed consistently high and relatively stable trust levels across stages. Participants with 5–8 years of experience displayed more variability, with trust peaking at the beginning and decreasing notably by the collision-avoidance stages. Conversely, those observers with over 8 years of experience started lower and exhibited a slight downward trend.

Finally, the age-based interaction highlights that participants over 40 years old exhibited relatively stable and higher trust scores (see Fig. 6 (c)), while those younger than 30 had more pronounced declines, particularly from Stages 2 to 4. Together, these interaction effects emphasise that trust is not only influenced by system actions but also by demographic characteristics.

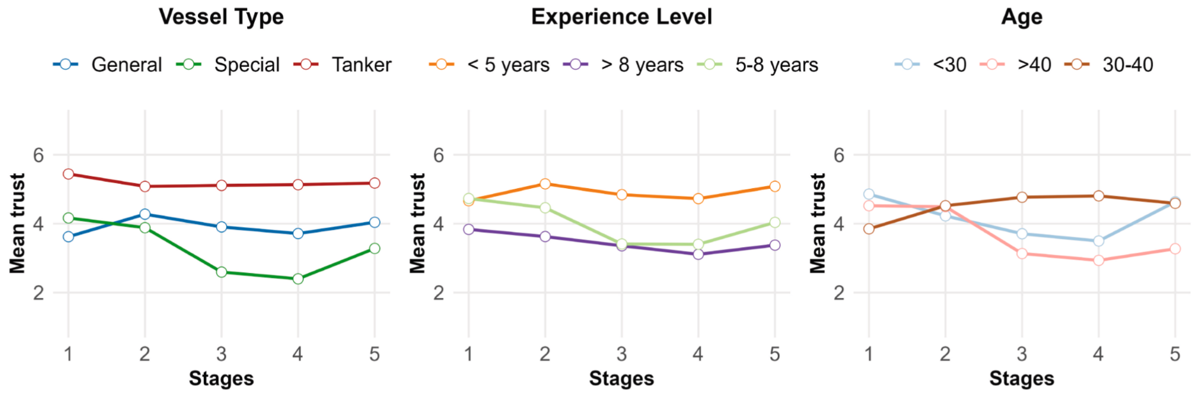
Five dimensions of trust

To gain insight into the key dimensions shaping observers' trust in MASS's navigation, we conducted a factor analysis on five trust-related metrics: *Dependability*, *Predictability*, *Anthropomorphism*, *Faith*, and *Safety*. Preliminary tests confirmed that the dataset was suitable for factor analysis, with a Kaiser-Meyer-Olkin (KMO) value of 0.843 (indicating sampling adequacy) and a significant Bartlett's Test of Sphericity was significant ($\chi^2 = 365.757$, $p < 0.001$). The factor analysis yielded a two-factor solution, explaining 88.16 % of the variance, indicating a stable structure in trust assessments (see Fig. 7). Factor 1 accounts for 67.6 % of the variance and includes *Dependability*, *Predictability*, *Anthropomorphism*, and *Faith*, while Factor 2 explains an additional 20.6 % and is represented solely by *Safety*. The extracted factors reveal that observers assess trust along two distinct dimensions: general System Competence and Situational Safety.

Specifically, the first factor, which we labelled "System Competence", aggregates four dimensions: Dependability, Predictability, Anthropomorphism, and Faith. As shown in Fig. 7, each of these dimensions has a strong loading on Factor 1. *Dependability* and *Predictability* capture the reliability and consistency of the MASS' navigation, while *Anthropomorphism* and *Faith* add human likeness and forward-looking trust, respectively. The second factor, labelled "Situational Safety", is defined exclusively by the safety-related dimension, which loads solely on this factor. Unlike the broad reliability-based attributes found in Factor 1, *Safety* reflects observers' perceptions of safety during collision avoidance.

Correlation analysis of trust dimensions and two related factors

Following the factor analysis, a correlation analysis was conducted to further investigate the relationships between the two key factors of trust and trust levels across different operational stages. This analysis aimed to understand how perceptions of trust evolve during the stages of navigation and how they correlate with the two identified trust factors.



(a) trust scores across 5 stages under different vessel types
(b) trust scores across 5 stages under different experience levels
(c) trust scores across 5 stages under different ages

Fig. 6. Trust dynamics across five navigation stages in relation to vessel type, experience level, and age.

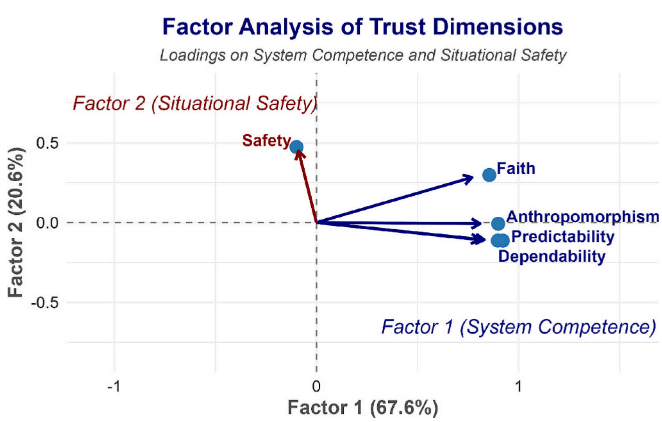


Fig. 7. The illustration of the factor analysis on five trust-related dimensions.

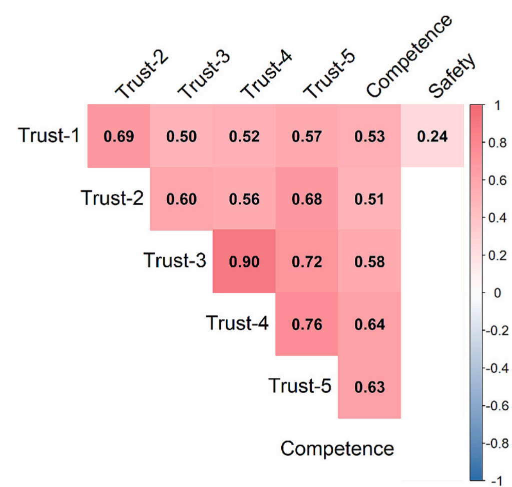


Fig. 8. Correlation matrix between trust scores across each stage and the two identified components.

Using Pearson’s correlation coefficients, we assessed the strength and direction of relationships between the five trust stages and the two factors identified in the factor analysis. Only significant correlations were visualised in the matrix, with non-significant cells left blank to emphasise meaningful associations. As illustrated in Fig. 8, The correlation matrix presents a series of moderate to strong positive correlations among trust scores across various stages. Additionally, trust scores between adjacent stages show the highest correlations, such as Trust1 and Trust2 (0.69) and Trust3 and Trust4 (0.90), indicating that trust levels evolve sequentially as participants progress through the stages.

System competence exhibited moderate positive correlations with trust scores across various stages (ranging from 0.51 to 0.64), underscoring the consistent influence of perceived competence on participants’ trust. In contrast, *Situational safety* displayed no significant correlations with the trust scores at stages other than the trust at stage 1. This result aligns with the earlier factor analysis, where *Situational safety* emerged as a distinct factor.

Building on the insights from our exploratory analysis, which highlighted key demographic and experimental conditions influencing trust, we propose a BN model for trust to capture these complex dynamics. This model formalises the relationships among System Competence, Situational Safety, stage-specific trust levels, and demographic and situational variables (strategies and timings), allowing us to quantify the influence of each factor on trust formation and development.

Trust model design

Bayesian network construction for trust

BN were selected as the trust modelling tool in this study due to their capability to represent complex probabilistic dependencies among multiple interacting variables while managing uncertainty. Compared to conventional regression models, BN provides a structured and interpretable approach for capturing conditional dependencies and sequential trust progression across navigation stages. This is particularly suitable for modelling trust in human-autonomy interaction contexts, such as MASS navigation, where trust evolves dynamically based on situational factors and observer characteristics.

Building on the insights from our exploratory analysis, which highlighted key demographic and experimental conditions influencing trust, we propose a BN model for trust to capture these complex dynamics. This model formalises the relationships among System Competence, Situational Safety, stage-specific trust levels, and demographic and situational variables (strategies and timings), providing a structured framework to quantify the influence of each factor on trust formation and development.

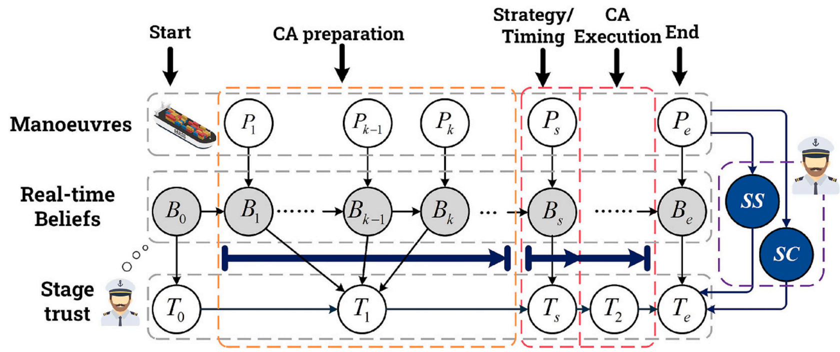


Fig. 9. Development of a human trust model with Bayesian Networks for MASS operation.

Node definition and network structure

The Bayesian network incorporates five sequential trust nodes, each representing trust at a specific stage, from *InitialTrust* to *FinalTrust*. This structure leverages the Markov property, as was considered in (Kok and Soh, 2020), where each trust stage depends solely on the trust level of its immediate predecessor. By adopting this assumption, the model focuses on the local dependencies in trust evolution, simplifying the structure while preserving the temporal dynamics of trust development. *FinalTrust* serves as the node that represents the cumulative confidence built throughout the CA process. It reflects how trust, as it propagates through the stages, aggregates into an overall assessment of the navigational performance of the autonomous system.

In addition to temporal dependency, trust varies among participants across various backgrounds, such as age, experience, and vessel types. Thus, this model integrates demographics that were identified as key factors, including age, experience, and vessel type, as parent nodes to InitialTrust, reflecting their role in shaping baseline trust levels. These factors account for inherent individual differences in trust propensity, as indicated by the exploratory findings. Furthermore, situational factors such as Strategies and Timings are introduced as parent nodes to Trust 3, representing the influence of CA decisions on trust in the decision-making stage. This structure ensures that the model captures both individual propensity and situational factors on trust transitions.

To capture the multidimensional evaluation of trust, the model incorporates two extracted components: *System competence* and *Situational Safety*. *System Competence* reflects perceptions of dependability, predictability, human likeness, and forward-looking beliefs, while *Situational Safety* focuses on safety evaluations during collision avoidance. These dimensions are linked directly to *FinalTrust*, representing their role in shaping the overall trust in the autonomous system. This framework lays the groundwork for further analysis, including diagnostic analysis informed by sensitivity analysis, predictive reasoning, and causal inference, to explore trust mechanisms in depth.

Fig. 9 illustrates the staged trust formation process of MASS in the CA process, showing the interaction between performance, real-time beliefs, and stage-specific trust across navigation phases. Trust evolves sequentially, starting with *initial* trust (T_0) and baseline beliefs (B_0), and progressing through key stages, including T_1 (routine navigation), T_s (strategy and timing decisions), T_2 (CA execution), and final trust T_e. At each stage, real-time beliefs (B_k) are updated dynamically based on ongoing system performance P_k, which directly shape staged trust. During CA execution, the system's manoeuvres (e.g., CA strategies and timing) are captured in performance nodes (P_s), which influence T_s via updated beliefs (B_s). Throughout the process, observer evaluations of System Competence (*SC*) and Situational Safety (*SS*) are integrated into final trust judgments. These two dimensions are critical to linking specific system performance to comprehensive trust evaluations at the final stage. This framework highlights the interplay of system performance, real-time beliefs, stage trust, and trust-related factors assessment in trust formation. Given the uniformity of vessel performance and the controlled nature of the experimental scenarios, performance variability was minimal. As such, the model excludes explicit *performance* nodes, focusing instead on Strategy and Timing as key situational factors of trust.

Parameter setting and model training

The constructed trust Bayesian network is shown in Fig. 10, where InitialTrust serves as the baseline trust level influenced by demographic factors, including age and vessel type, which were derived from maritime industry reports,[1,2] For example, age distributions (below 30: 16 %, 30–40: 29 %, above 40: 55 %) and vessel type (General: 63 %, Tanker: 13 %, and Special-purpose ships: 25 %). For factors lacking statistical support, such as Strategy and Timing, prior probabilities were estimated based on domain expertise. For instance, left-turns (25 %) and right-turns (75 %) were assigned probabilities reflecting standard maritime practices under COLREGs, while collision-avoidance timing was set as standard (70 %), early (15 %) and imminent (15 %). Additionally, System Competence and Situational Safety were discretised into low, medium, and high categories using tertile thresholds (0.33 and 0.66) derived from factor analysis scores, while trust ratings (1–7) were similarly classified into low (1–2), medium (3–5), and high (6–7). The prior probabilities of other nodes and conditional probabilities were calculated by using the trust data collected from our survey through the Genie software.

Application

Following the construction of the TBN model, its utility was evaluated through targeted applications. These included **diagnostic analysis** informed by sensitivity insights and **predictive reasoning**. Diagnostic analysis, built on sensitivity analysis methods, aims to identify the most influential factors contributing to a specific observed outcome.. Predictive inference estimates future trust levels based on current conditions, aiding in proactive management.

Diagnostic analysis

To evaluate the robustness and identify critical determinants of the trust model, we conducted a diagnostic analysis informed by sensitivity insights targeting the Trust5 = high outcome. A 30 % parameter spread, reflecting realistic variability in parameters, was implemented to simulate realistic uncertainties, visualising results using a tornado diagram (see Fig. 11), where the top ten bars represent the factors contributing most significantly to the variability of the outcome.

As shown in Fig. 11, the tornado diagram highlights the diagnostic

[1] https://www.statista.com/statistics/264024/number-of-merchant-ships-worldwide-by-type/.

[2] https://www.gov.uk/government/statistical-data-sets/seafarer-statistics-sfr#certificated-officers-and-trainees-sfr02.

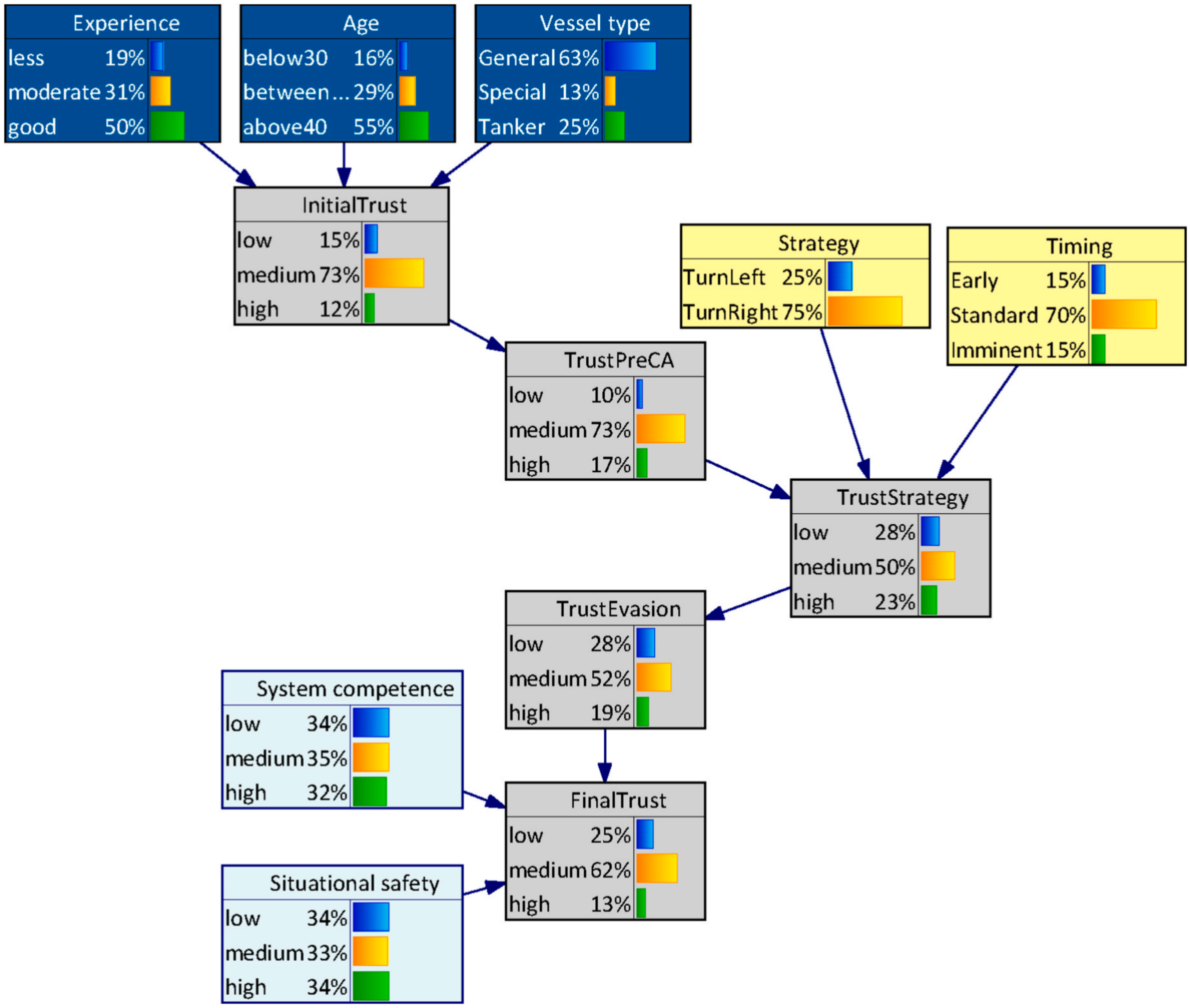


Fig. 10. Trust model design for autonomous decision-making of MASS in CA scenarios.

results of Trust5 = high to variations in key parameters, demonstrating how trust outcomes respond to changes in the TBN. Competence = high exhibits the most significant positive influence, aligning with its direct pathway to *FinalTrust* and underscoring its central role in trust formation. Sequential trust stages, such as Trust4 = high | Trust3 = high, reveal cascading effects, emphasising the importance of consistent trust-building across stages. Together, the insights emphasise the interplay between System Competence and sequential trust evolution, offering actionable guidance for enhancing user trust in autonomous navigation systems.

In TBN, Trust 3 represents a critical stage where trust is influenced by the prior trust level, that is, TrustPreCA, and situational factors (e.g., Strategies, Timings). This node is important to explore because it indirectly impacts *FinalTrust*, as identified in Fig. 11 (The second most important impact factor: Trust4 = high|Trust3 = high). In addition, it is the key stage in the whole process at which the Strategy and Timing were imposed. Thus, the diagnostic analysis for Trust3 = high was conducted further, as shown in Fig. 12. Specifically, the analysis reveals that Trust2 = medium, conditional on Trust1 = medium, exerts the strongest influence, with a steep negative derivative (−0.207), indicating that small changes in Trust2 greatly impact Trust3. Similarly, the interaction between Timing=Standard and Strategy = TurnRight demonstrates a marked influence on Trust3 = medium, evidenced by its contribution and derivative (−0.133). Notably, the direct influence of Timing=Standard (ranked 5th) compared to its interaction with Strategy (ranked 2nd and 3rd) highlights the compounding effect of navigation strategies on trust. This aligns with the finding that Strategy = TurnRight combined with a higher trust level in Trust2 contributes positively to Trust3 = high (derivative: +0.276). Furthermore, while other factors also show the impact on Trust3, such as vessel type = General, their effects are weaker, underscoring the dominance of imminent variables such as situational factors over demographics.

Similarly, a diagnostic analysis on Trust1 = high was also conducted, as shown in Fig. 13. The results reveal that Vessel Type exhibits the

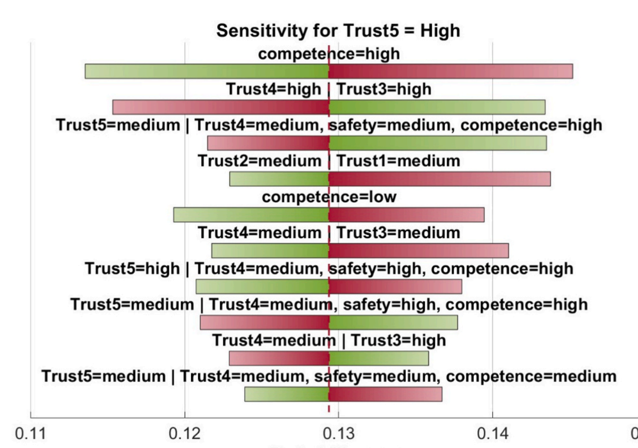


Fig. 11. Diagnostic analysis visualisation results for Trust5(Final Trust) = high.

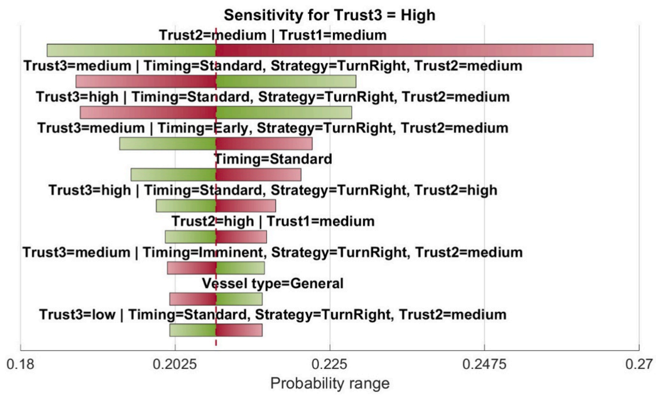


Fig. 12. Diagnostic analysis visualisation results for Trust3(TrustStrategy) = high.

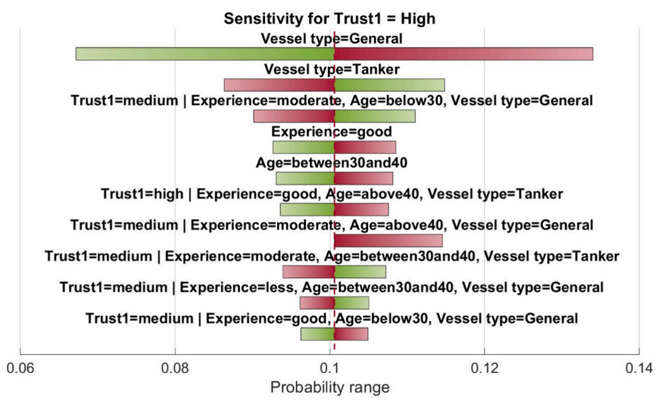


Fig. 13. Diagnostic analysis visualisation results for Trust1(InitialTrust) = high.

strongest influence on Trust1, particularly for general vessels showing a negative relationship (derivative: −0.170) and tanker vessels with a positive influence (derivative: +0.206), indicating a higher trust dependency on vessel types. Other demographic factors such as Experience and Age demonstrate moderate but substantial effects, with experienced participants (rated as "good") and those aged 30–40 exhibiting negative impacts on Trust1 = high. Conversely, specific combinations of demographic features (e.g., good experience and vessel type "tanker") highlight positive influence, reflecting that senior, experienced personnel on takers enhance *InitialTrust*. This analysis underscores the importance of tailoring strategies to specific observer profiles to foster trust in autonomous systems from the outset.

Predictive reasoning

Following the diagnostic analysis, we conducted predictive reasoning to estimate the trust dynamics under the variations in Strategies and Timings, particularly focusing on critical trust nodes, such as *FinalTrust* and *TrustEvasion*, see Fig. 14. As shown in Fig. 14 (a), medium trust consistently dominates, with the Right & Early strategy achieving the highest proportion (60 %). High trust levels are, although relatively low, peak in Right & Standard (28 %), indicating its effectiveness in maintaining trust during the evasive stage. Conversely, low trust is most prevalent in Left & Standard (33 %), suggesting its potential drawbacks in trust-sensitive scenarios. Similarly, the FinalTrust subplot, as shown in Fig. 14 (b), shows medium trust as the dominant outcome across all strategies, with Right & Early achieving the highest proportion (64 %) and Left & Standard again exhibiting higher low trust levels (27 %).

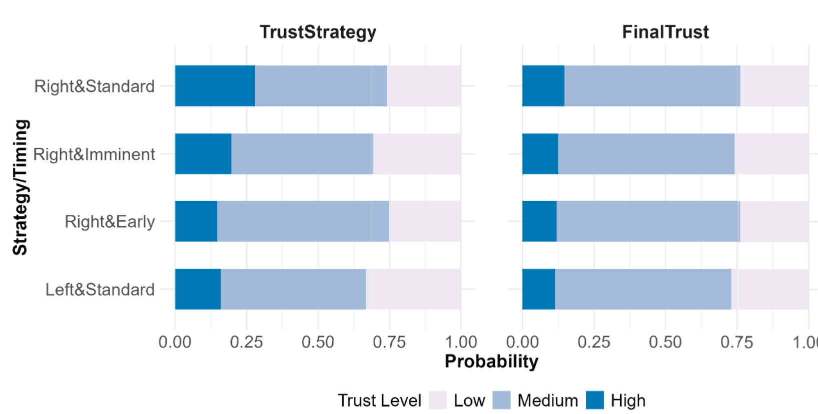
These findings underscore the diagnostic finding of trust outcomes to operational strategies, highlighting Right & Standard, Right & Early, and Right & Imminent as favourable strategy combinations for sustaining trust during the entire CA process.

Discussion

Interpretation of results

Trust was measured using post-scenario evaluations collected via Qualtrics, where participants rated their trust after observing specific collision avoidance manoeuvres. The use of simulated navigation videos embedded within *Qualtrics* ensured that participants evaluated the autonomous system's performance in controlled, consistent scenarios, capturing trust fluctuations across distinct navigation stages. Furthermore, the analysis, conducted using LMM, uncovered trust dynamics across navigation stages.

Aligned with H1: Firstly, consistent with H1, trust in the MASS fluctuated throughout the CA process. Participants' trust varies significantly across several stages (e.g., TrustPreCA vs TrustStrategy) but partially recovered during the final stage, see Fig. 4. This fluctuation reflects increased scrutiny during high-stakes manoeuvres and a gradual convergence towards a calibrated level of trust as participants gained a deeper understanding of the system. However, the final trust levels did not return to their initial levels, suggesting residual caution or incomplete trust recovery even after successful task completion. Secondly, trust levels exhibit slight increases during the early stages (Trust1 to



(a) Trust Evasion under Different Collision Avoidance Strategies

(b) Final Trust under Different Collision Avoidance Strategies

Fig. 14. Predictive reasoning on trust evasion and final trust under different CA Strategies.

Trust2), reflecting trust accumulation, but a shapely decrease in Trust3 and Trust4, underscoring the asymmetric nature of trust formation versus erosion, followed by partial recovery at the final stage (Trust5). The initial slight increase may result from the system's adherence to stable navigation practices and predictable behaviour. The abrupt decline likely corresponds to participants' heightened scrutiny during strategies/timings selection and execution stages, where system limitations or perceived inefficiencies become more evident. Trust recovery at the final stage suggests an accumulation effect, where the overall performance in earlier stages is synthesised into a final trust judgment. This pattern aligns with trust accumulation, typically requiring consistent system performance over time, while dissipation can occur rapidly due to a single negative event.

Aligned with H2: Furthermore, trust in the Right&Standard scenario differs significantly from the Left & Standard scenario, as shown in Fig. 5, suggesting participants' preference for manoeuvres that align more closely with COLREGs. This result, aligning with H2, may turn out that in CA scenarios, where a vessel is approaching from her starboard side, right-turn strategies may have been perceived as more consistent with standard maritime practices to accumulate trust, while left-turn strategies might have been interpreted as riskier or less conventional to dissipate trust.

Aligned with H3: Finally, aligning with H3, while proactive responses aligned with standard timings were associated with higher trust levels, actions that were "too early" or "too late" demonstrated suboptimal outcomes, see Fig. 5. The findings imply that MASS systems must balance evasion strategies and proper timings, avoiding evasions that are either too proactive or overly reactive.

Overall, these two factors reveal that observers differentiate between general System Competence and Situational Safety when forming trust in autonomous navigation systems. This insight emphasises the need for MASS designs to address both Competence and Safety to ensure reliability and promote trust in dynamic navigational environments.

In terms of demographic factors consideration, the inclusion of participants with diverse professional backgrounds aimed to ensure the representativeness of trust dynamics across various groups. Thiis diversity allowed to identify overall trend in trust evoluation while also capturing the variability that emerges when demographic factors interact with other factors. The results indicate while the main effects analysis revealed that trust dynamics were primarily influenced by navigational stages and conditions, interaction effects highlighted subtle differences based on experience level, vessel type, and age during specific CA stages, as shown in Fig. 6. These differences were not the primary focus of this study but provide supplemetary insights into how trust responses may vary in certain CA scenarios. Such insights highlight the need for context-specific considerations when evaluating trust in MASS navigation in CA scenarios.

Regarding the dimensional structure of trust, trust was found to encompass two overarching dimensions: System Competence and Situational Safety. The linkage between System Competence and Situational Safety and *FinalTrust* demonstrates the multidimensional nature of trust. This finding highlights that observers evaluate trust both as a comprehensive judgment of the system's competence and as a context-specific assessment of safety. Additionally, *System competence* exhibited moderate positive correlations with trust scores across various stages (ranging from 0.51 to 0.64), underscoring the consistent influence of perceived competence on participants' trust. This result suggests that observers' perceptions of the MASS's navigational reliability, human likeness, and forward-looking beliefs contribute continuously to their trust across all stages, indicating their foundational role in trust formation. In contrast, situational safety was primarily linked to InitialTrust. Its influence on subsequent trust stages was limited. This may reflect the controlled nature of the experimental design, in which participants were implicitly assured of the system's safety. In other words, in this context, *Safety* might become a "given" in participants' minds, leading them to assume that the MASS will handle high-risk scenarios adequately. As a result, Safety ratings might remain stable across different conditions, especially if no unexpected behaviours challenge this expectation. However, this does not imply that situational safety is irrelevant in real-world applications. Instead, it suggests that observers' perceptions of safety are formed early and remain stable unless disrupted by unexpected system failures or high-risk scenarios.

With respect to TBN, this model captures the staged progression of trust in MASS, integrating temporal dynamics, demographics, and situational factors. This structured approach is essential for understanding how trust evolves and identifying the determinants of trust-building at different stages of the CA process. Firstly, the sequential trust nodes represent a staged process of trust evaluation from *InitialTrust* to *FinalTrust*. The Markov property simplifies the model by assuming that each stage depends primarily on the previous one, which is consistent with the exploratory analysis showing strong correlations between consecutive trust ratings, see Fig. 8. Secondly, baselined trust levels (InitialTrust) are influenced by demographic variables, such as vessel type, age, and experience (see Fig. 13 and Fig. 6(a)).

Focusing on the results of diagnostic analysis informed by sensitivity insights, two aspects of insights can be drawn. (1) The tornado diagram for Trust5 = high (Fig. 11) indicates that System Competence exerts the most significant positive influence on *FinalTrust*. It underscores that perceptions of dependability, predictability, human likeness, and forward-looking beliefs of the autonomous system in the entire CA process are critical for building overall trust. (2) The cascading influence of earlier trust stages on later outcomes (e.g., Trust4 = high | Trust3 = high) emphasises the cumulative nature of trust (Fig. 11). The significant impact of *TrustStrategy* (Trust3) on *FinalTrust* highlights the critical role of decision-making strategies and timings in the trust pathway. Furthermore, TrustStrategy (Trust3) was found to be influenced not only by situational factors (e.g., strategy and timing) but also by the trust level in the preceding stage (TrustPreCA). This sequential dependency supports the hypothesis that trust evolves progressively, with earlier stages laying the foundation for subsequent evaluations. The findings support the need for consistent trust-building throughout all stages of interaction.

Finally, the following key takeaways can be derived regarding the results of predictive reasoning: (1) strategies involving Right & Early, Right & Imminent, and Right & Standard manoeuvres consistently achieve higher levels of trust compared to Left strategies, as shown in Fig. 14, also aligning with the hypothesis of H3. (2) Despite variations during evasive actions, trust partially stabilises at the *FinalTrust* stage. This indicates that the system's overall performance, which affects the system competence of the autonomous system, can mitigate earlier fluctuations, reinforcing the importance of holistic trust-building efforts.

Implications of findings

Overall, the results have the following two aspects of implications for the design and operation of autonomous navigation systems.

(1) Prioritising competence in system design: System Competence was underscored, comprising reliability, predictability, anthropomorphism, and forward-looking decision-making, as the most critical factor influencing observer overall trust in the entire CA process. MASS systems should prioritise performance consistency and predictability, especially in CA scenarios. To achieve this, developers must enhance the transparency of system behaviour by incorporating real-time feedback mechanisms that clarify decision rationales, particularly during unconventional manoeuvres such as left-turn strategies. Additionally, to maintain trust consistently, MASS systems must focus on early-stage performance to prevent dissipation that could propagate through later evaluations.
(2) Optimising evasion strategies and timing: The study highlights the importance of proper evasive strategy and timing. While

proactive responses are generally associated with higher trust levels, actions that are too early or too delayed can dissipate observer trust. To address this, MASS systems should incorporate adaptive algorithms that optimise the timing of evasive manoeuvres with compliance with regulations like COLREGs. Furthermore, autonomous systems should focus on transparency, particularly in explaining the decision logic in scenarios where deviations from observer expectations (e.g., delayed or unconventional manoeuvres) occur. As suggested by (Song et al., 2024a), observer trust in autonomous navigational decisions can be strengthened when the regulations are involved in the decision-making mechanism, which can improve the system's transparency.

In terms of comparison with prior research, the findings align with previous studies on trust in automation, particularly the dynamic nature of trust, accumulation and dissipation (Alhaji et al., 2023), and its dependence on system performance (Xu and Dudek 2015a). In the maritime domain, this study, which is different from (Poornikoo et al., 2024), investigates observer trust in the autonomous decision-making system of MASS across several stages in a CA process instead of real-time measurement. However, this study expands the understanding of trust in autonomous systems by introducing the dual dimensions of competence and situational safety, providing an in-depth understanding of trust in the autonomous system of MASS in an MWTS.

Limitations

Despite its contributions, this study has several limitations that warrant further investigation. The use of a simulator-based experimental setup, while providing a controlled and realistic maritime environment through the ship manoeuvring simulator, cannot fully replicate the complexity of real-world MASS navigation. Factors such as variable sea states, multi-ship encounters, and communication delays in actual operations were not incorporated, which may affect ecological validity.

Additionally, the sample size (N = 26), although diverse in maritime experience, limits the generalisability of findings across the broader seafaring population. The participant pool did not specific ally include officers with experience on passenger-carrying vessels, whose heightened safety responsibilities might influence trust perceptions differently. Future studies should aim to include this demographic to broaden the applicability of the findings.

Furthermore, this study focuses solely on observational aspects of human supervision in MASS operations. While this mirrors real-world RCC scenarios, it does not capture trust dynamics involving direct human intervention, which may need further investigation.

The TBN model relies on Markovian assumptions, simplifying trust progression to local dependencies and potentially overlooking long-term influences. Moreover, trust was assessed post-scenario, which may not fully capture real-time fluctuations during critical events. Furthermore, participants' relatively stable evaluations of situational safety may reflect the controlled nature of the experiment, where baseline expectations shaped their perceptions.

Conclusions & future research

This study investigated the dynamics of observer trust in MASS during CA scenarios, combining quantitative trust measurement, exploratory analysis using LMM, and predictive reasoning via the proposed TBN model.

Trust was measured through post-scenario evaluations collected via Qualtrics, allowing participants to rate their trust in MASS after observing simulated navigation videos. These measurements captured stage-specific fluctuations, which were analysed using LMM to identify key patterns: slight and gradual trust accumulation during routine navigation and sharp dissipation during the CA strategies and timings selection and execution stages. Trust at the final stage, that is, overall trust, is partially recovered, underscoring the cumulative influence of prior stages. Trust dynamics varied significantly by demographic factors, such as experience and vessel type. Moreover, left-turn strategies were associated with lower trust compared to right-turn strategies, reflecting observer preferences for COLREGs-compliant evasion strategies. Factor analysis identified two trust dimensions, including System Competence and Situational Safety, with System Competence strongly correlating with trust across all stages. The Markov-like stage correlations further supported the sequential nature of trust evolution.

Building on these findings, the TBN model quantified trust dynamics, highlighting the dominant role of System Competence in shaping final trust and the cascading influence of intermediate stages. Diagnostic analysis informed by sensitivity analysis emphasised the critical importance of decision-making strategies and timely actions, while predictive reasoning demonstrated the positive impact of proactive right-turn manoeuvres. These insights provide actionable guidance for designing MASS systems that align with observer expectations, improve transparency, and optimise CA strategies.

Future research should validate these findings in real-world maritime contexts and extend trust modelling to encompass variable environmental conditions. Moreover, integrating physiological data and real-time monitoring tools could offer deeper insights into trust fluctuations during dynamic navigation tasks. Expanding participant diversity, particularly involving officers with experience on passenger-carrying vessels and including both observational and intervention-based supervision, will further enhance the validity and applicability of trust models in MASS operations. These efforts will contribute to developing autonomous navigation systems that are not only technically robust but also aligned with human supervisory expectations in complex maritime environments.

CRediT authorship contribution statement

Rongxin Song: Writing – original draft, Validation, Software, Methodology, Conceptualization. **Eleonora Papadimitriou:** Writing – review & editing, Supervision, Conceptualization. **Rudy R. Negenborn:** Writing – review & editing, Supervision. **Pieter van Gelder:** Writing – review & editing, Supervision, Resources, Project administration, Conceptualization.

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.

Acknowledgements

This study is the result of the research project funded by the China Scholarship Council under grant 202106950002 and supported by Researchlab Autonomous Shipping (RAS) of Delft University of Technology. Additionally, we are deeply grateful to the Wuhan University of Technology for providing access to simulator-based technical support, which was essential to the completion of this research.

Data availability

Data will be made available on request.

References

Alhaji, B., Prilla, M., Rausch, A., 2023. Robot collaboration and model reliance based on its trust in human-robot interaction. Lecture Notes in Computer Science (including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). https://doi.org/10.1007/978-3-031-42283-6_2.

Alhaji, B., Prilla, M., Rausch, A., 2021. Trust dynamics and verbal assurances in human robot physical collaboration. Front. Artif. Intell. https://doi.org/10.3389/frai.2021.703504.
Alsos, O.A., Saghafian, M., Veitch, E., Petermann, F.-M., Sitompul, T.A., Park, J., Papachristos, E., Eide, E., Breivik, M., Smogeli, Ø., 2024. Lessons learned from the trial operation of an autonomous urban passenger ferry. Transp. Res. Interdiscip. Perspect. 26. https://doi.org/10.1016/j.trip.2024.101142.
Bakdi, A., Vanem, E., 2022. Fullest COLREGs evaluation using fuzzy logic for collaborative decision-making analysis of autonomous ships in complex situations. IEEE Trans. Intell. Transp. Syst. 23 (10), 18433–18445. https://doi.org/10.1109/TITS.2022.3151826.
Basu, C., Singhal, M., 2016. Trust dynamics in human autonomous vehicle interaction: A review of trust models. AAAI Spring Symposium - Technical Report 85–91.
Chang, C.H., Wijeratne, I.B., Kontovas, C., Yang, Z., 2024. COLREG and MASS: Analytical review to identify research trends and gaps in the Development of Autonomous Collision Avoidance. Ocean Eng. 302, 117652. https://doi.org/10.1016/J.OCEANENG.2024.117652.
Chauhan, H., Pakbaz, A., Jang, Y., Jeong, I., 2024. Analyzing trust dynamics in human–robot collaboration through psychophysiological responses in an immersive virtual construction environment. J. Comput. Civ. Eng. 38 (4), 04024017. https://doi.org/10.1061/JCCEE5.CPENG-5692/ASSET/DCFB6559-ABF5-44BB-8607-6DE926C0E544/ASSETS/IMAGES/LARGE/FIGURE5.JPG.
Chen, M., Nikolaidis, S., Soh, H., Hsu, D., Srinivasa, S., 2018. Planning with trust for human-robot collaboration, ACM/IEEE International Conference on Human-Robot. Interaction 307–315. https://doi.org/10.1145/3171221.3171264.
Chen, M., Soh, H., Hsu, D., Nikolaidis, S., Srinivasa, S., 2020. Trust-aware decision making for human-robot collaboration: Model learning and planning. ACM Trans. Hum. Robot Interact. 9, 2. https://doi.org/10.1145/3359616.
Fooladi Mahani, M., Jiang, L., Wang, Y., 2021. A Bayesian trust inference model for human-multi-robot teams. Int. J. Soc. Robot. 13 (8), 1951–1965. https://doi.org/10.1007/s12369-020-00705-1.
Gil, M., 2021. A concept of critical safety area applicable for an obstacle-avoidance process for manned and autonomous ships. Reliab. Eng. Syst. Saf. 214, 107806. https://doi.org/10.1016/J.RESS.2021.107806.
Gregor, A.W.H., Allison, R.S., Heffner, K., 2023. Exploring the Impact of Immersion on Situational Awareness and Trust in Remotely Monitored Maritime Autonomous Surface Ships, in: OCEANS 2023 - Limerick, OCEANS Limerick 2023. doi:10.1109/OCEANSLimerick52467.2023.10244249.
Guo, Y., Yang, X.J., 2021. Modeling and predicting trust dynamics in human–robot teaming: A bayesian inference approach. Int. J. Soc. Robot. 13 (8), 1899–1909. https://doi.org/10.1007/s12369-020-00703-3.
Hancock, P.A., Billings, D.R., Schaefer, K.E., Chen, J.Y.C., De Visser, E.J., Parasuraman, R., 2011. A meta-analysis of factors affecting trust in human-robot interaction. Hum. Factors 53 (5), 517–527. https://doi.org/10.1177/0018720811417254/ASSET/IMAGES/LARGE/10.1177_0018720811417254-FIG1.JPEG.
Hoff, K.A., Bashir, M., 2015. Trust in automation: Integrating empirical evidence on factors that influence trust. Hum. Factors 57 (3), 407–434. https://doi.org/10.1177/0018720814547570.
Hopko, S.K., Mehta, R.K., Pagilla, P.R., 2023. Physiological and perceptual consequences of trust in collaborative robots: An empirical investigation of human and robot factors. Appl. Ergon. 106, 103863. https://doi.org/10.1016/J.APERGO.2022.103863.
Hou, M., Ho, G., Dunwoody, D., 2021. IMPACTS: a trust model for human-autonomy teaming. Human-Intelligent Systems Integration 2021 3:2 3 2, 79–97. doi:10.1007/S42454-020-00023-X.
Huang, Y., 2019. Supporting human-machine interaction in ship collision avoidance systems. Delft University of Technology, Delft, Netherlands.
Jackson, K.M., Sharp, W.H., Shaw, T.H., 2022. Modeling the impacts of positive interaction frequency on subjective trust in an autonomous agent: A linear mixed model approach. https://doi.org/10.1177/1071181322661448.
Kirkpatrick, J., Hahn, E.N., Haufler, A.J., 2017. Trust and Human–Robot Interactions. Robot Ethics 2.0: From Autonomous Cars to Artificial Intelligence 142–156. doi:10.1093/OSO/9780190652951.003.0010.
Kok, B.C., Soh, H., 2020. Trust in robots: Challenges and opportunities. Curr. Robot. Rep. 1 (4), 297–309. https://doi.org/10.1007/s43154-020-00029-y.
Krausman, A., Neubauer, C., Forster, D., Lakhmani, S., Baker, A.L., Fitzhugh, S.M., Gremillion, G., Wright, J.L., Metcalfe, J.S., Schaefer, K.E., 2022. Trust measurement in human-autonomy teams: Development of a conceptual toolkit. ACM Trans. Hum. Robot. Interact. 11, 3. https://doi.org/10.1145/3530874/ASSET/E3BCC44A-FEF8-4AF4-9BA7-F1994A5A2136/ASSETS/GRAPHIC/THRI-2020-0088-F08.JPG.
Lee, J., Feng, J., Humt, M., Müller, M.G., Triebel, R., 2021. Trust your robots! Predictive uncertainty estimation of neural networks with sparse gaussian processes. Proceedings of Machine Learning Research.
Lee, J.D., See, K.A., 2004. Trust in automation: Designing for appropriate reliance. https://doi.org/10.1518/hfes.46.1.50_30392.
Luo, R., Weng, Y., Jayakumar, P., Brudnak, M.J., Paul, V., Desaraju, V.R., Stein, J.L., Ersal, T., Yang, X.J., 2024. Real-time workload estimation using eye tracking: A bayesian inference approach. Int. J. Hum. Comput. Interact. 40 (15), 4042–4057. https://doi.org/10.1080/10447318.2023.2205274.
Lynch, K., Plant, K.L., Roberts, A.P.J., Banks, V.A., Taunton, D., 2023. Investigating decision-making in the operation of Maritime Autonomous Surface Ships using the Schema World Action Research Method. In: ACM International Conference Proceeding Series. https://doi.org/10.1145/3597512.3597521.
Lynch, K.M., Banks, V.A., Roberts, A.P.J., Radcliffe, S., Plant, K.L., 2024. What factors may influence decision-making in the operation of Maritime autonomous surface ships? A systematic review. Theor. Issues Ergon. Sci. 25 (1), 98–142. https://doi.org/10.1080/1463922X.2022.2152900.
Mallam, S.C., Nazir, S., Sharma, A., 2020. The human element in future Maritime Operations–perceived impact of autonomous shipping. Ergonomics 63 (3), 334–345. https://doi.org/10.1080/00140139.2019.1659995.
Malle, B.F., Ullman, D., 2020. A multidimensional conception and measure of human-robot trust, in: Trust in Human-Robot Interaction. doi:10.1016/B978-0-12-819472-0.00001-0.
Mayer, R.C., Davis, J.H., Schoorman, F.D., 1995. An integrative model of organizational trust. Acad. Manag. Rev. 20 (3), 709. https://doi.org/10.2307/258792.
Misas, J.P., Hopcraft, R., Tam, K., 2022. Future of maritime autonomy: Cybersecurity, trust and mariner's situational awareness. In: Proceedings of the International Ship Control Systems Symposium. https://doi.org/10.24868/10703.
Negenborn, R.R., Goerlandt, F., Johansen, T.A., Slaets, P., Valdez Banda, O.A., Vanelslander, T., Ventikos, N.P., 2023. Autonomous ships are on the horizon: here's what we need to know. Nature 2023 615:7950 615 7950, 30–33. doi:10.1038/d41586-023-00557-5.
Park, E., Jenkins, Q., Jiang, X., 2008. Measuring trust of human operators in new generation rescue robots. In: Proceedings of the JFPS International Symposium on Fluid Power. https://doi.org/10.5739/ISFP.2008.489.
Poornikoo, M., Gyldensten, W., Vesin, B., Øvergård, K.I., 2024. Trust in automation (TiA): Simulation model, and empirical findings in supervisory control of maritime autonomous surface ships (MASS). Int. J. Hum. Comput. Interact. 1–28. https://doi.org/10.1080/10447318.2024.2399439.
Roy, S.K., Singh, G., Sadeque, S., Gruner, R.L., 2024. Customer experience quality with social robots: Does trust matter? Technol. Forecast Soc. Change 198. https://doi.org/10.1016/j.techfore.2023.123032.
Sadrfaridpour, B., Saeidi, H., Burke, J., Madathil, K., Wang, Y., 2016. Modeling and control of trust in human-robot collaborative manufacturing. In: Robust Intelligence and Trust in Autonomous Systems. https://doi.org/10.1007/978-1-4899-7668-0_7.
Sheng, S., Pakdamanian, E., Han, K., Kim, B.G., Tiwari, P., Kim, I., Feng, L., 2019. A case study of trust on autonomous driving. 2019. IEEE Intelli. Transp. Syst. Conf. ITSC 2019, 4368–4373. https://doi.org/10.1109/ITSC.2019.8917251.
Soh, H., Xie, Y., Chen, M., Hsu, D., 2020. Multi-task trust transfer for human–robot interaction. Int. J. Rob. Res. 39 (2–3), 233–249. https://doi.org/10.1177/0278364919866905.
Song, R., Papadimitriou, E., Negenborn, R.R., van Gelder, P., 2024a. Integrating situation-aware knowledge maps and dynamic window approach for safe path planning by maritime autonomous surface ships. Ocean Eng. 311, 118882. https://doi.org/10.1016/J.OCEANENG.2024.118882.
Song, R., Papadimitriou, E., Negenborn, R.R., van Gelder, P., 2024b. Safety and efficiency of human-MASS interactions: towards an integrated framework. J. Mar. Eng. Technol. https://doi.org/10.1080/20464177.2024.2414959.
van de Merwe, K., Mallam, S., Nazir, S., Engelhardtsen, Ø., 2024. Supporting human supervision in autonomous collision avoidance through agent transparency. Saf. Sci. 169. https://doi.org/10.1016/j.ssci.2023.106329.
Wang, M., Hussein, A., Rojas, R.F., Shafi, K., Abbass, H.A., 2018. EEG-based neural correlates of trust in human-autonomy interaction. In: Proceedings of the 2018 IEEE Symposium Series on Computational Intelligence. https://doi.org/10.1109/SSCI.2018.8628649.
Wróbel, K., Gil, M., Huang, Y., Wawruch, R., 2022. The Vagueness of COLREG versus Collision Avoidance Techniques—A Discussion on the Current State and Future Challenges Concerning the Operation of Autonomous Ships. Sustainability 2022, Vol. 14, Page 16516 14 24, 16516. doi:10.3390/SU142416516.
Wu, B., Hu, B., Lin, H., 2017. Toward efficient manufacturing systems: A trust based human robot collaboration, in: Proceedings of the American Control Conference. pp. 1536–1541. doi:10.23919/ACC.2017.7963171.
Xu, A., Dudek, G., 2015. OPTIMo: Online Probabilistic Trust Inference Model for Asymmetric Human-Robot Collaborations. ACM/IEEE International Conference on Human-Robot Interaction 2015-March, 221–228. doi:10.1145/2696454.2696492.
Xu, C., Zhang, C., Zhou, Y., Wang, Z., Lu, P., He, B., 2024. Trust recognition in human-robot cooperation using EEG. Proc. - IEEE Int. Conf. Robot. Automation 7827–7833. https://doi.org/10.1109/ICRA57147.2024.10610156.
Yagoda, R.E., Gillan, D.J., 2012. You want me to trust a ROBOT? The development of a human-robot interaction trust scale. Int. J. Soc. Robot. 4 (3), 235–248. https://doi.org/10.1007/S12369-012-0144-0/MEDIAOBJECTS/12369_2012_144_MOESM1_ESM.PDF.
Zarzycki, F., Gil, M., Montewka, J., Szłapczyński, R., Szłapczyńska, J., 2025. Declarative ship arenas under favourable conditions. Ocean Eng. 316, 119927. https://doi.org/10.1016/J.OCEANENG.2024.119927.
Zheng, H., Smereka, J.M., Mikulski, D., Wang, Y., 2023. Bayesian optimization based trust model for human multi-robot collaborative motion tasks in offroad environments. Int. J. Soc. Robot. 15 (7), 1181–1201. https://doi.org/10.1007/S12369-023-01011-2/FIGURES/11.