

Article

Towards an Emotion-Aware Metaverse: A Human-Centric Shipboard Fire Drill Simulator

Musaab H. Hamed-Ahmed ¹, Diego Ramil-López ^{1,2}, Paula Fraga-Lamas ^{1,2,3,*}
and Tiago M. Fernández-Caramés ^{1,2,3}

¹ Department of Computer Engineering, Faculty of Computer Science, Universidade da Coruña, 15071 A Coruña, Spain; musaab.ahmed@udc.es (M.H.H.-A.); diego.ramil.lopez@udc.es (D.R.-L.); tiago.fernandez@udc.es (T.M.F.-C.)

² Centro Mixto de Investigación UDC-Navantia, Universidade da Coruña, Edificio de Batallones, s/n, 15403 Ferrol, Spain

³ Centro de Investigación CITIC, Universidade da Coruña, 15071 A Coruña, Spain

* Correspondence: paula.fraga@udc.es; Tel.: +34-981167000

Abstract: Traditional Extended Reality (XR) and Metaverse applications focus heavily on User Experience (UX) but often overlook the role of emotions in user interaction. This article addresses that gap by presenting an emotion-aware Metaverse application: a Virtual Reality (VR) fire drill simulator for shipboard emergency training. The simulator detects emotions in real time, assessing trainees' responses under stress to improve learning outcomes. Its architecture incorporates eye-tracking and facial expression analysis via Meta Quest Pro headsets. Two experimental phases were conducted. The first revealed issues like poor navigation and lack of visual guidance. These insights led to an improved second version with a refined User Interface (UI), a real-time task tracker and clearer visual cues. The obtained results showed that the included design improvements can reduce task completion times between 14.18% and 32.72%. Emotional feedback varied, suggesting a need for more immersive elements. Overall, this article provides useful guidelines for creating the next generation of emotion-aware Metaverse applications.



Academic Editor: Juan Gabriel Avina-Cervantes

Received: 27 April 2025

Revised: 7 June 2025

Accepted: 14 June 2025

Published: 17 June 2025

Citation: Hamed-Ahmed, M.H.; Ramil-López, D.; Fraga-Lamas, P.; Fernández-Caramés, T.M. Towards an Emotion-Aware Metaverse: A Human-Centric Shipboard Fire Drill Simulator. *Technologies* **2025**, *13*, 253. <https://doi.org/10.3390/technologies13060253>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: Metaverse; Industrial Metaverse; emotions; Extended Reality; fire drill; maritime safety

1. Introduction

The Metaverse has emerged as a transformative concept, integrating XR technologies to create immersive and interactive environments [1]. Although the Metaverse was initially conceptualized in science fiction [2], it is now being adopted across industries, including education, healthcare or entertainment [3–5]. One of its most promising applications is related to the field of training, especially in safety and emergency preparedness, where immersive simulations can provide effective and scalable training solutions [6,7].

Specifically, maritime safety is a critical domain where fire emergencies pose significant risks to crew members, cargo and vessel integrity. According to the International Convention for the Safety of Life at Sea (SOLAS), regular fire drills are mandatory to ensure that crews are adequately trained to respond to onboard fires [8,9]. However, traditional fire drills are often logistically challenging, resource-intensive and difficult to personalize for individual crew members. XR-based fire drill simulations offer an alternative by providing realistic, cost-effective and repeatable training in a controlled environment.

A crucial aspect of emergency response is emotional resilience, which indicates the ability to remain calm and to make informed decisions under pressure. However, traditional

XR and Metaverse-oriented applications, although they can analyze the user inputs in terms of performance, they do not usually consider users' emotions. Thus, emotion-aware XR training enhances traditional simulations by integrating emotion-detection technologies, such as eye tracking and facial expression analysis, to assess and to adapt training scenarios in real time. By monitoring stress levels, cognitive load and decision-making behaviors, the system can personalize training to improve both technical proficiency and psychological preparedness. For instance, emergency-specific or mission-critical applications such as fire drills are particularly well suited, as they require realistic exposure to high-risk scenarios that evoke stress, urgency and the need for rapid decision-making conditions that are critical for preparing trainees to effectively respond in real-world emergencies. Metaverse and XR technologies are especially effective in this context, as they provide immersive and interactive environments capable of evoking the emotional and cognitive responses needed for effective training.

This article presents the design and evaluation of an emotion-aware VR fire drill simulator for shipboard emergencies, whose main contributions are the following:

- A holistic review is provided on the integration of XR technologies, emotion recognition methods and performance analytics to create emotion-aware Metaverse applications.
- It is described what to the knowledge of the authors is the first emotion-aware maritime safety training Metaverse application.
- A detailed evaluation is presented with groups of real final users (naval engineers) that assessed the developed solution in terms of performance and emotional engagement.

Specifically, this article is focused on addressing the following research questions:

- RQ1: How can emotion awareness be integrated into XR and Metaverse applications to improve user interaction, particularly in training simulations?
- RQ2: What is the current gap in traditional XR and Metaverse applications regarding the role of emotions in user experience?
- RQ3: What architectural components are necessary for developing an emotion-aware Metaverse application, specifically for real-time emotion detection?
- RQ4: How can eye-tracking and facial expression analysis be incorporated into XR headsets (e.g., Meta Quest Pro) to facilitate emotion detection in a Metaverse training environment?
- RQ5: Can trainees' performance be improved in a VR fire drill simulator by adding design elements and UI refinements?
- RQ6: Can the integration of real-time emotion detection in a VR fire drill simulator provide valuable insights into trainees' responses under stress?

The rest of this article is structured as follows. Section 2 reviews the state of the art and the background work in relation to the Metaverse, fire drill regulations, and the development of emotion-based XR solutions, as well as the most relevant hardware and software for detecting and monitoring emotions in Metaverse applications. Section 3 describes the design of the proposed emotion-aware fire drill simulator, whose implementation is detailed in Section 4. Section 5 is dedicated to the performed experiments. Finally, Section 6 details the main key findings and Section 8 is devoted to the conclusions.

2. State of the Art

This section begins with a detailed explanation of the concept of the Metaverse, followed by an overview of fire drill regulations to establish the regulatory context. It then reviews current applications of the Metaverse within the maritime industry, highlighting relevant use cases and advancements. The discussion transitions into the role of emotion

detection in XR, including both the theoretical foundations and practical techniques for detecting emotions using XR technologies. Building on this, the section explores emotion-aware applications within the Metaverse. Finally, it examines the hardware and software commonly employed for emotion detection in Metaverse-based applications.

2.1. About the Metaverse

The concept of “Metaverse” was initially introduced in 1992 as a continuous virtual realm seamlessly integrated into everyday life [10]. The term “Metaverse” derives from the Greek prefix “meta”, meaning “beyond”, and the root “-verse”, from “universe”, thus defining the Metaverse as a reality that transcends the physical universe. As of writing, no standardized definition has been established [11–13]. This ambiguity allows industry leaders to shape the concept in alignment with their perspectives and strategic objectives [14,15]. Nevertheless, despite the lack of consensus, the recognition of the Metaverse potential across industries underscores its broad and diverse opportunities [16].

The distinction between the Metaverse and XR environments further complicates its definition. Technologies such as Internet of Things (IoT) based home automation systems, which facilitate XR interactions [17], exhibit characteristics similar to those of the Metaverse but that are typically classified as Cyber-Physical Systems (CPSs) [18] rather than direct Metaverse innovations.

From a technical perspective, the Metaverse can be described as a massively scaled interoperable network of real-time rendered 3D virtual environments [19]. Users engage with the Metaverse synchronously and persistently, experiencing a continuous sense of presence while maintaining data continuity across identity, history, entitlements, objects, communications, and transactions. The concept transcends any single entity or industry, with various virtual worlds tailored to interests such as sports, entertainment, art, or commerce [20].

In practice, as of writing, multiple Metaverses coexist [21], forming collections of interconnected virtual worlds governed by hierarchical structures analogous to the Internet complex architecture and the geopolitical organization of the physical world. The terminology for these collections is still under discussion [22], with proposals including “Multiverse” [11] and “MetaGalaxy” [23].

2.2. Fire Drill Regulations

Fire drills are essential for ship management and are overseen by the ship master [8]. Due to the severe threat of fire accidents to lives, cargo and financial interests, regular fire drills are mandated [24].

SOLAS, established in 1974 and subsequently amended, stands as the cornerstone of maritime safety and environmental protection regulations. SOLAS encompasses eleven chapters, each addressing specific aspects of maritime safety and operation [24]. Specific to fire safety, SOLAS mandates various regulations addressing fire detection, extinguishing, Personal Protective Equipment (PPE), evacuation procedures, and dangerous materials handling [25]. Key regulations include the following:

1. The availability of visual and audible alarms (Regulation II-2/7 and III 6.4.2).
2. Availability of proper evacuation through escape routes, and muster areas (Regulation II-2/12).
3. Response to fire emergency by specifying the necessary firefighting equipment, providing crew training in firefighting methods, and establishing protocols for containing and extinguishing fires (Regulation II/2 2.1.7, 5, 7.5.1, 15.2.1.1, 15.2.3, 16.2, 18.8 and III 35).

4. Regular fire drills so the crew acquaints themselves with the evacuation procedures (IMO 2009) (Regulation II-2/15.2.2 and III/19.3).
5. Availability of proper signage and navigation lighting on the escape routes (Regulation II/13.3.2.5).

Specifically, SOLAS regulation III/19.3 indicates that fire drills must be conducted monthly for each crew member, simulating realistic fire scenarios, with a focus on vulnerable ship areas [25]. Standardized fire drill scenarios typically involve several steps [26], which are illustrated in Figure 1.

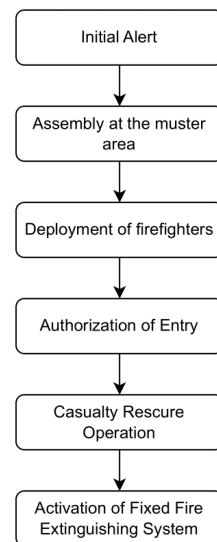


Figure 1. Steps performed for fire drills onboard ships.

A fire drill unfolds in a structured sequence, beginning with the sounding of fire alarms upon the initial alert of a fire incident. Crew members swiftly assemble at predetermined muster stations for attendance verification. Then, designated firefighters, guided by the master's instructions, prepare to confront the fire. Upon authorization from the master, firefighting personnel engage in suppression activities while prioritizing casualty rescue operations. In parallel, reserve personnel stand ready to support firefighting endeavors. Should initial firefighting attempts prove inadequate, the activation of the fixed fire extinguishing system becomes imperative, serving as a crucial backup measure in containing the escalating situation [26].

There are evaluation criteria for each step of the fire drill, assessing factors like time, performance and task completion [8]. Following drills, the ship master evaluates crew performance, providing corrective actions as needed to enhance preparedness [8].

2.3. Metaverse Applications for Maritime Industries

While XR technology has demonstrated effectiveness across industries, its integration into maritime operations remains limited [27]. For instance, previous research has proposed three primary uses for the use of XR in maritime operations [28]: accident prevention, accident response and accident analysis. Accident prevention involves risk assessment and management during routine operations and emergencies. Accident response focuses on evacuation processes during emergencies, while accident analysis translates documented accidents into virtual environments for analysis.

Augmented Reality (AR) finds extensive use for maritime construction, aiding workshop workers during outfitting [29,30] and assembly processes [31]. AR applications can integrate real-time sensor measurements with technology platforms like Unity, thus facilitating streamlined information retrieval processes [32,33].

Regarding VR, there are some relevant solutions for non-maritime industries, such as the serious game presented in [34], which was developed for fire preparedness training in a CAVE (Cave Automatic Virtual Environment) system targeting elementary school teachers. Other authors investigated the effects of immersion on both conceptual and procedural learning in various fire scenarios (e.g., hospital bed fire, a vehicle engine fire, an electrical fire, and a lab bench fire) [35]. In the maritime sector, an interesting solution is presented in [36], where a simulation-based learning platform is presented, which is accessible to all Norwegian maritime education and training institutions.

VR is commonly employed in training simulators due to its cost-effectiveness and environmental friendliness [37]. Applications include spray painting training [38], fire-fighting simulations [39], safety training [40,41] or dynamic risk assessment in mooring operations [42]. Multi-role [43] and multi-player [44] simulations have also been developed for lifesaving and evacuation strategies aboard passenger ships. In addition, it is worth noting that prominent maritime organizations have adopted VR technology for training and educational initiatives, spanning safety protocols, inspection procedures and onboard operations [40].

2.4. Emotion Detection for XR

2.4.1. Detecting Emotions with XR Technologies

Emotion detection can become a fundamental component for XR applications where user engagement, stress response, and decision making are critical. In immersive training environments, such as a Metaverse-based fire drill simulator, monitoring emotional states provides valuable insights into user behavior under pressure. By integrating emotion recognition mechanisms, these systems can dynamically adjust training scenarios and enhance realism, learning outcomes and overall preparedness for real-world emergencies.

The detection of emotional states in immersive environments relies on physiological and behavioral data analysis, encompassing facial expressions, eye movement tracking, heart rate variability and galvanic skin response [45]. These bio signals offer real-time feedback on user engagement and affective states, enabling systems to deliver more adaptive and responsive experiences. However, emotion recognition in virtual settings presents several challenges, including inter-individual and cross-cultural variations in emotional expression and sensor accuracy limitations [46]. Differences in physiological and psychological factors influence the manifestation and interpretation of emotions, necessitating the use of diverse datasets to enhance system reliability. Moreover, the precision of emotion detection depends on hardware quality, including cameras, eye tracking systems and wearable sensors, which may introduce constraints in dynamic environments.

Other studies, such as Zhang et al. [47], focused on VR pedagogies and the role of textual cues within vocational education. Their findings indicate that vicarious VR significantly enhances immediate knowledge acquisition, although it has little effect on long-term retention and transfer learning. While textual cues generally did not impact learning outcomes, they did improve immediate knowledge acquisition in direct manipulation VR; however, they were found to be unnecessary in vicarious VR contexts. In related work, Vontzalidis et al. [48] explored the use of spatial audio cues. Meanwhile, Jing et al. [49] identified a lack of standardized representation of emotion in collaborative VR.

Finally, it is also worth mentioning that some authors have previously studied the use of XR technologies for Empathic Computing, which includes not only emotion recognition techniques, but also context awareness, adaptive interaction, feedback collection mechanisms, or user behavior modeling [50]. As can be guessed, Empathic Computing systems can be really complex, being able to detect emotions through a wide variety of mechanisms, including spoken words, facial expressions, behavioral patterns or physiological responses.

2.4.2. XR Emotion Detection Techniques

Emotion recognition in immersive environments is achieved through multiple techniques, including:

- Facial expression analysis: tracking facial movements enables the inference of emotional states based on expressions such as frowning, smiling or eyebrow motion [51]. Currently, only a few head-mounted displays incorporate inward-facing sensors to capture subtle facial cues, improving avatar realism and scenario adaptation [52].
- Eye tracking: the analysis of gaze patterns, pupil dilation and blink rates provides insights into stress levels, cognitive load, and emotional engagement [46]. These data are particularly valuable for assessing decision-making processes in high-pressure simulations.
- Physiological signal monitoring: wearable sensors are able to measure physiological indicators such as heart rate, skin conductivity, or temperature fluctuations to detect emotional arousal [53]. These metrics help when evaluating stress and engagement levels in training scenarios.
- Voice analysis: variations in speech tone, pitch and rhythm can reveal emotional states such as urgency, hesitation or confidence, making this method particularly useful for assessing verbal responses in interactive training exercises [46].

2.5. Emotion-Aware Applications on the Metaverse

The Metaverse offers substantial potential for the development of emotion-driven applications, which enhance user interaction and engagement within virtual environments. By integrating emotional responsiveness into these digital spaces, platforms can create more immersive and adaptive experiences that respond to users' emotional states. Such applications are proving to be beneficial across a range of sectors, including industrial training, healthcare and collaborative workspaces, where they improve the realism of simulations, support safety protocols, and enable a higher degree of personalization. The mentioned sectors may differ in the context and emotional intensity of their scenario (e.g., healthcare), but they share key similarities in high-stakes decision making, time-sensitive operations and the need for precision under pressure. Thus, such domains operate within mission-critical environments where emotional regulation and situational awareness are essential for effective performance. As such, adopting a broad perspective enriches the overall discussion and highlights transferable insights across domains.

For instance, in the healthcare domain, platforms like XRHealth have incorporated emotion-sensitive avatars to support therapeutic interventions [54]. An example of a use case is the treatment of post-traumatic stress disorder, for which XRHealth adjusts exposure therapy based on the user's emotional feedback, employing facial expression analysis to modify the intensity of the session. This personalized approach fosters a controlled yet dynamic environment that adapts to the patient's emotional needs, facilitating recovery in ways that traditional therapy may not be able to achieve. In addition, virtual environments can be utilized for rehabilitation, where patients recovering from injuries or surgeries interact with scenarios that are adjusted according to their emotional responses, providing both physical and emotional support. Moreover, it is worth indicating that Artificial Intelligence (AI) driven virtual environments are revolutionizing mental well-being interventions. For example, the Merrytopia project [55] combines natural language processing and Machine Learning to analyze users' journal entries and emotional states. Such a platform generates personalized recommendations and immersive mindfulness scenes, dynamically adjusting interactions through real-time sentiment analysis. This creates a positive feedback loop where an AI-driven chatbot reinforces constructive aspects of users' experiences, offering scalable solutions for stress and anxiety management. Thus, the developed system

demonstrates how adaptive virtual environments can complement traditional therapeutic methods by tailoring support to individual emotional needs.

Beyond healthcare, the Metaverse is also being leveraged to enhance emotional intelligence and interpersonal skills. An example is a study that introduced a Metaverse-based English teaching solution that makes use of emotion-based analysis methods to personalize the learning experience [56]. This approach creates an immersive and interactive teaching environment, enabling educators to adjust lessons based on students' emotional states, thereby increasing engagement and improving learning outcomes.

In the educational field, XR-based systems can harness the multimodal data to optimize learning processes. For instance, previous research has made use of eye-tracking, facial expression analysis and electrodermal activity sensors to demonstrate that reduced blink rates and sustained visual attention on task-relevant content correlate strongly with improved knowledge retention [57]. These findings enable the design of adaptive systems that respond to cognitive load indicators in real time, thus enabling one to adjust the complexity of procedural training scenarios.

2.6. Hardware and Software for Detecting Emotions in Metaverse Applications

Effective emotion detection in industrial Metaverse environments relies on advanced hardware and intelligent software systems. Modern AR and Mixed Reality (MR) headsets, such as Meta Quest Pro [58] or Microsoft HoloLens [59], integrate high-resolution cameras (up to 13 MP) and sophisticated eye tracking technologies (e.g., Vive Focus 3 Eye Tracker [60]). Such headsets are able to monitor gaze direction and pupil dilation, which are key indicators of cognitive states like focus and fatigue. In addition, facial tracking devices, such as the HTC Vive Facial Tracker [61], enable the real-time detection of microexpressions and lip movements, helping to infer emotions like happiness, surprise or stress.

Specialized accessories further enhance emotion-recognition capabilities. Haptic gloves, such as Teslaglove [62], incorporate physiological sensors that measure heart rate and blood oxygen levels, providing insight into stress and anxiety. Full body tracking devices, including Sony Mocopi [63] and HaritoraX [64], analyze posture and movement patterns, which can be linked to emotional states (e.g., a slouched posture indicating frustration). Omnidirectional treadmills like Virtuix Omni One [65] track locomotion behaviors, correlating movement speed and patterns with agitation or calmness. By using advanced haptic feedback systems, such as bHaptics TactSuit X16 [66], immediate tactile responses can be provided, including chest vibrations during high-stress situations, which can help in user guidance and decision making. Moreover, emotion adaptive visual interfaces enhance situational awareness by adjusting lighting and color schemes in immersive environments to maintain focus and to reduce cognitive overload.

3. Design of an Emotion-Aware Metaverse Application: A Fire Drill Simulator

This section outlines the main objective of the simulator, presents the system architecture and details the design of the emergency scenario.

3.1. Objective of the Application

A fire drill simulator helps carry out training for fire emergencies that might occur on board vessels. Its objective is to enhance trainees' decision making and risk assessment capabilities by simulating various fire scenarios and allowing the rehearsal of key emergency procedures in a controlled manner.

Adding the component of emotional control within simulated procedures enhances realism, considering real fire emergency situations would certainly induce emotions like

fear and stress, which could affect performance. With the incorporation of emotional recognition, the system not only assesses procedural competencies but also emotional resilience, thus providing personalized feedback and enhancing preparedness for high-pressure situations.

It is important to note that the devised VR application was designed to act like a Metaverse where users can join to train their shipboard fire drill skills and engage with persistent content, including their identity, history, entitlements, objects, communications, and transactions. Multiple trainees can make use of the developed application; however, to simplify the emotion-detection analyses described in this article, for the experiments performed in Section 5, the multiple participating users interacted with the application in a non-concurrent way, like in a traditional immersive VR experience.

3.2. System Architecture

Figure 2 depicts the overall communications architecture of the proposed system. Such an architecture illustrates how users interact with the developed Metaverse applications: the users (i.e., Meta-Operators) wear Metaverse-ready XR devices that act as gateways to enter the different Metaverses provided by the available MetaGalaxies. Some of such MetaGalaxies are included as an example, but there are many more, which can be tailored to specific fields (e.g., industry, entertainment, gaming). Moreover, it is important to note that this article essentially deals with the lower components of the proposed architecture (i.e., with the interactions and emotions of the users through a Metaverse-Ready XR device), so the communications with other MetaGalaxies are beyond the scope of this article.

Despite the definition of Metaverse depicted in Figure 2, it is worth indicating that, in the literature, other authors provide different visions of the concept of Metaverse [67–70], since, as it was previously indicated in Section 2.1, there is still not a standardized definition. Thus, although some authors consider that there is only one Metaverse that includes multiple virtual worlds (which was the original concept of Metaverse proposed in 1992), in this article, it is considered that the current situation is actually closer to Figure 2, where each MetaGalaxy contains a network of interconnected Metaverses thanks to a framework that allows for sharing protocols or ecosystems. It is worth pointing out that, in such a latter scenario, moving from one Metaverse to another requires a unified digital identity system (e.g., a universal login, a blockchain-based digital passport), which enables maintaining users' avatars, assets, and preferences across platforms. Moreover, in a MetaGalaxy-based ecosystem, interoperability standards are crucial, since they allow different virtual worlds to support common formats for avatars, 3D models and user interactions. Using the previously mentioned underlying technologies, transitions between Metaverses could occur through in-world portals, menu-based navigation, or hyperlink-like connections. Therefore, seamless asset portability and synchronized user data would ensure continuity, while governance systems would manage access and would enforce local rules. This vision is still an ongoing effort that depends on broad cross-platform collaboration and the adoption of open and standardized protocols [71,72].

In addition, it is worth indicating that the selected Metaverse-ready XR devices are essential in the proposed architecture, since they are the ones responsible for detecting emotions. Thus, such devices should make use of the necessary hardware and software to detect emotions, and also implement the required Application Programming Interfaces (APIs) to interact with the rest of the system to provide feedback to the users or to adapt the developed application to the trainees' emotions. In order to deliver fast reactions to the user, such tasks should ideally be performed locally (i.e., on the XR device), but applications that require heavy data processing can make use of external high-performance Edge Computing devices like Cloudlets [73].

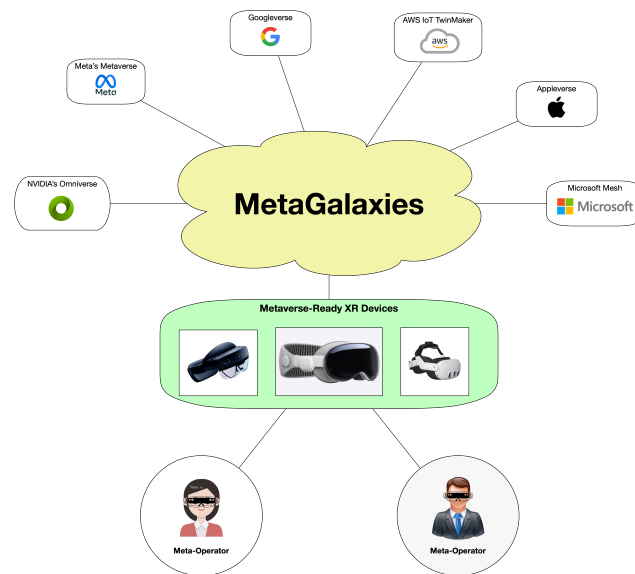


Figure 2. General communications architecture.

Figure 3 illustrates through a sequence diagram the main component interactions of the flow related to the on-device part of the architecture. As it can be observed, the trainee first executes the Metaverse application, which starts to monitor their emotions. Once an emotion is detected, the Metaverse application is notified, which derives into two potential actions: (1) to provide feedback to the user (e.g., through visual cues or text instructions); and/or (2) to adapt the application to the detected emotions (e.g., to increase visual stimuli to generate the expected emotions).

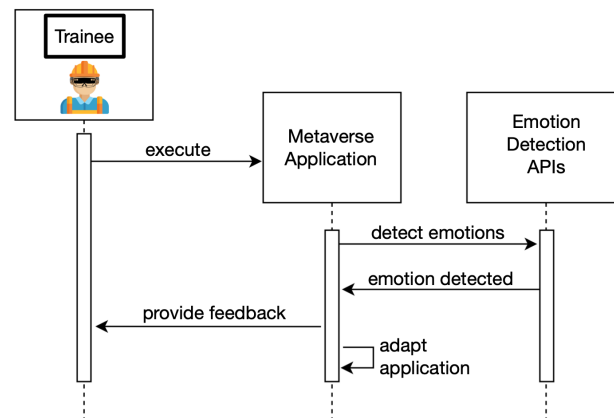


Figure 3. Sequence diagram on the overall interactions of the on-device flow.

3.3. Design of the Emergency Scenario

Since crafting scenarios for every conceivable ship area proves impractical, previous research [74,75] has underscored the heightened likelihood of fire occurrences in specific ship zones, notably the engine room and the accommodation area, particularly the galley. These areas are prone to fire incidents due to factors such as the presence of machinery and electrical equipment in the engine room, along with the abundance of oil and fuel, which pose ignition risks. Similarly, electrical faults in the galley can swiftly escalate into fire outbreaks. Consequently, the engine room and galley were selected as focal points for scenario development, aligning with empirical evidence of fire prevalence in these ship sections.

When crafting the scenario, it was considered that the main objective was to create an immersive training experience for crew members to simulate fire emergencies and to

evaluate their response capabilities, behavior, and risk analysis skills. Insights gathered on shipboard fire drill scenarios [8,26] inform the design process, supplemented by guidance from a fire drill training and operation booklet [76]. The scenario, developed with the previously described critical elements in mind, unfolds as follows:

1. The Meta-Operators to be trained navigate freely within the virtual environment, situated either in the engine room or the galley.
2. Their task involves promptly identifying the fire's location, signaled by auditory cues like burning sounds or visual cues such as flames.
3. Upon detection, participants inform the ship master about the fire location and size, and activate the fire alarm system.
4. They then assess the fire's severity to determine whether it is controllable and extinguishable, or whether it poses an imminent threat.
5. Finally, participants evacuate to the designated muster area.

Following such stages, the simulator takes the trainees within different levels where the intensity of the scenario increases. This concept draws inspiration from the notion of gamification [77–79], which follows a design approach of applying game elements such as engagement, challenge, and structured progression to non-game contexts in order to promote motivation and sustained interaction. In the implementation presented in this article, gamification is included through the different stages within the application, each presenting diverse fire intensity levels and fire eruption locations. Thus, the proposed multi-level approach serves as an effective means of evaluating the trainees' performance during the whole training process.

4. Implementation of the Metaverse Application

This section provides a comprehensive overview of the implementation. It begins by describing the hardware and software components utilized in the development of the fire drill simulator. Second, it outlines the definition of the four scenario levels design for the simulator. This is followed by a detailed specification of the virtual environment. Finally, the section presents the features related to emotion detection, detailing how emotional responses are captured.

4.1. Hardware and Software

Among the potential headsets to be used, Meta Quest Pro was selected thanks to its ability to monitor natural facial expressions and to track the user eyes. Specifically, the Quest Pro features advanced eye-tracking technologies use inward-facing sensors to monitor and to interpret eye movements in real time.

Regarding the selected software, Unity Editor version 2022.3.15f1 was used with the Universal Rendering Pipeline (URP). The XR Interaction Toolkit facilitated the creation of immersive interactions within the application. The OpenXR Software Development Kit (SDK) (1.034) was employed for communicating the Meta Quest Pro with the developed application, ensuring cross-platform compatibility and future-proofing [80]. The Face Tracking API from the Meta Movement SDK was used, since it allows developers to map detected expressions onto these characters for natural interactions. Thus, such an API converts movements into expressions based on the Facial Action Coding System (FACS).

In addition, Blender was used alongside Unity for specific segments of the virtual environment due to its robust feature set and VR readiness [81].

4.2. Scenario Levels

To construct the devised scenario, four main scenes were defined based on two key factors: the location (e.g., galley, engine room) and whether the fire is extinguishable or not. Each scene corresponds to a specific level in the simulator, as illustrated in Figure 4.

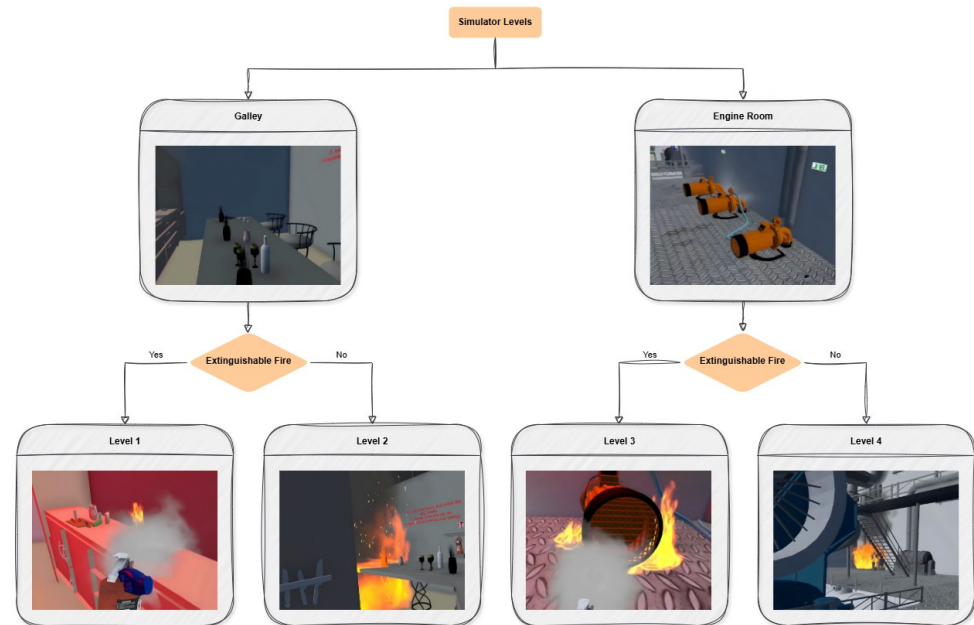


Figure 4. Levels of the fire-drill simulator.

Initially, in the first level, trainees are introduced to a scenario where a small fire erupts in the galley area, offering a manageable firefighting situation to acquaint them with basic fire management techniques. In this level, guidance text is provided to the trainee to perform the required tasks (as it is illustrated in Figure 5). As they advance to the second level, the intensity of the fire escalates, presenting a heightened challenge with an inextinguishable blaze, thereby testing the trainees' adaptability and decision making under pressure. Moving on to the third level, a new challenge emerges as a fire breaks out in the engine room, requiring trainees to navigate through the scenario without explicit instructions, relying instead on established safety protocols accessible via the UI instruction menu. Despite the increased complexity, the fire remains extinguishable at this stage, fostering critical thinking and procedural adherence. Finally, the fourth level presents the maximum difficulty, featuring a substantial fire outbreak in the engine room that rapidly intensifies. Here, trainees face an inextinguishable inferno, presenting the ultimate test of their firefighting prowess and resilience in the face of escalating adversity.

Considering the previously described level design, it can be easily guessed that it would be straightforward to incorporate emotion detection. For instance, fire intensity can be adapted to reach a desired level of emotions in order to deliver a realistic experience. There are many adaptations that can be performed on the devised fire drill scenario to consider emotional feedback, but the analysis and evaluation of such modifications go beyond this article, which, in the next sections, focuses on studying the technical feasibility and performance of the underlying technologies.

To assess the performance of the trainees, a script was created to log the critical interactions of the trainees with the required game objects. Thus, every interaction is recorded together with a timestamp and with the event, whether it was grabbing an object or activating it. This record later allows for estimating the user workload during the training.



Figure 5. Example of guidance text in Level 1.

4.3. Virtual Environment

The implementation of the virtual scenario involved importing a cargo ship model sourced and constructing the ship interior cabin. Blender, in conjunction with the Home-Builder add-on, facilitated the design, inspired by a tween-decker general cargo ship arrangement [82]. For instance, Figures 6 and 7 show, respectively, the created galley area and spawn room. Moreover, Figure 8 shows the engine room.

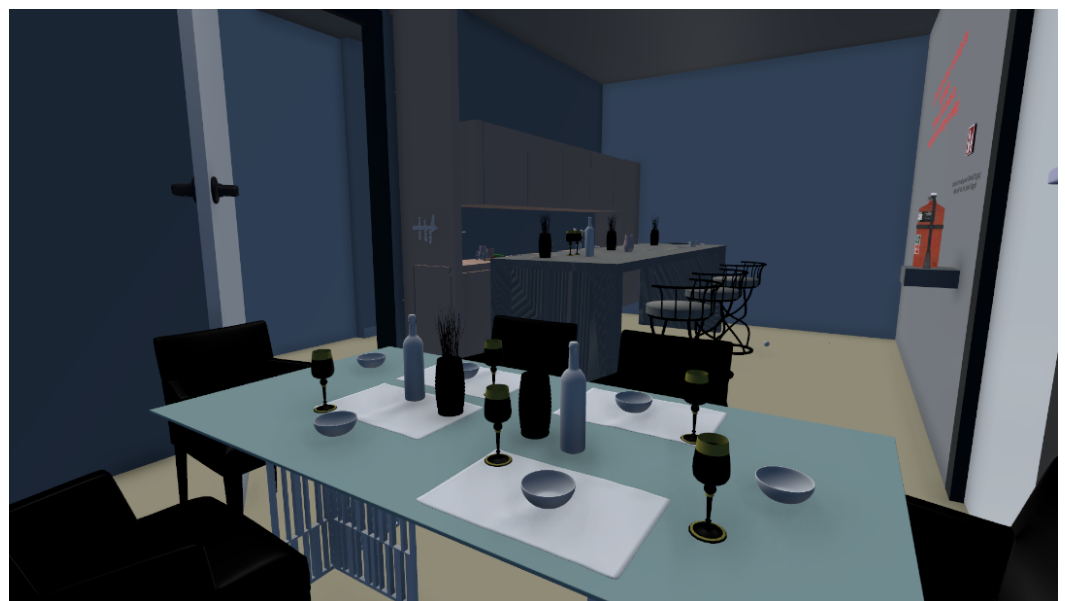


Figure 6. Galley area of the ship.

Lighting was applied using Unity's point light game object for optimal visibility. Collider components were implemented to enhance interaction realism. Interactive elements such as fire extinguishers, alarms, and emergency phones were included. Simulated fire and smoke effects were integrated by using Unity particle system. For example, Figure 9 shows the implemented fire extinguisher when it was used by a trainee to put off a simulated fire.



Figure 7. Spawn room of the application.



Figure 8. Engine room of the ship.



Figure 9. Moment when a trainee performed the fire extinguishing process.

Finally, a Sky High Dynamic Range Imaging (HDRI) scene and an ocean shader from Unity were included to add realism to the scenario (these can be observed in Figure 10).

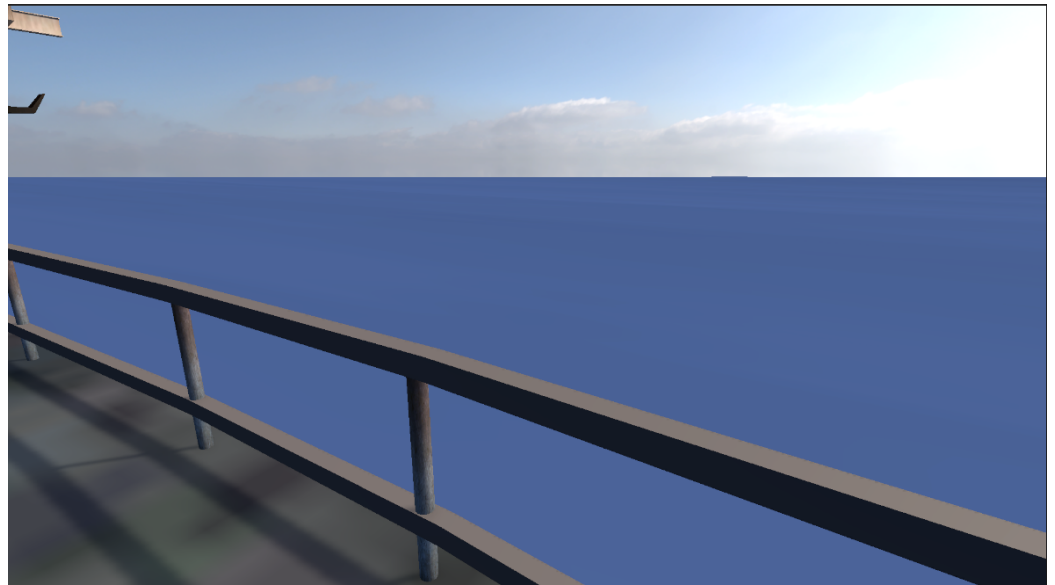


Figure 10. Ocean shader and sky scene added to the simulator.

4.4. Emotion Detection

Emotion detection focused on two separate features: eye tracking and natural facial expressions. To assess the workload of the trainees, the following actions were performed:

- An eye tracker script was implemented using the Meta Movement SDK, with its built-in Eye Gaze script. Moreover, another script was created to ensure the application of the appropriate ray cast and proper naming of the game objects.
- Facial expressions were logged with the Meta Movement SDK and the built-in Face Expressions API. Such an API translated the camera and the inward-sensors into blendshapes that monitored the logged expression with its strength through the respective weighting of each expression. In addition, a second script processes the weighted expressions to determine, through a rule-based logic, the detected emotions (e.g., happiness, sadness, surprise). For instance, rule-based logic refers to the use of predefined mapping based on FACS, which identifies emotions based on combinations of facial muscle movements and that link specific facial expressions to emotional states. The system monitors trainees' facial data in real time and compares it to the expected emotional responses defined for each scenario. Accuracy is calculated by measuring the match between expected and detected emotions.
- Lastly, the readings from both the eye tracking and the face expressions were stored simultaneously in a dataset for each level of the simulator. In this way, the logging mechanism enables the simulator to precisely determine when a trainee is directing their gaze toward a specific object in the virtual environment. By synchronizing this gaze data with the real-time detection of facial expressions, the system can then associate the emotional response with the corresponding stimulus. For example, if a trainee looks at a simulated fire and a fear-related facial expression is detected at that moment, the system logs this association. This approach allows for a contextual analysis of emotional responses, linking them directly to interactions with specific elements within the scenario.

5. Experiments

This section describes the experimental setup and presents a comprehensive evaluation of the performance of the simulator. Key performance metrics include the time required to complete each level, the number and sequence of gazed objects and a comparison of user performance with and without visual cues. In addition, the section evaluates the effectiveness of emotion detection, outlining the experimental conditions, reporting detection accuracy and analyzing the emotions identified for each user across the different levels.

5.1. Description of the Experiments

The experiments were conducted in two distinct phases, each evaluating a different version of the application:

- The first experiment was conducted with an earlier version of the application that only included textual cues to help the trainees through the tasks, but no visual cues were provided on how to perform the tasks (neither a map nor navigation clues to explore the virtual map). Ten naval engineers (8 men, 2 women) from the Erasmus Mundus Joint Masters Degree Sustainable Ship and Shipping 4.0 (SEAS 4.0) (imparted by the University of Napoli Federico II, the University of A Coruña and the University of Zagreb) participated in the first experiment. The participants represented a diverse international cohort, with testers from Pakistan, Nigeria, Azerbaijan, Sudan, Turkey, Cameroon, Colombia, Iran, Russia and Peru. Their ages ranged from 20 to 30 years, and they exhibited varying levels of prior experience in three key areas relevant to the study: fire drills, virtual reality (VR) and video gaming. Before the tests began, each participant was instructed on how to use the controllers, the purpose of the simulator, and how to conduct the fire drill. For this first experiment, the completion time to finish each level was measured. The reason to focus on completion time is because, in a shipboard drill, training performance and effectiveness are directly linked to such a time: a trainee needs to follow very specific steps in a fast manner to extinguish a fire, so their performance will be inversely proportional to the amount of time required to complete the necessary tasks (i.e., the longer the completion time, the lower the trainee performance). In addition, as a trainee learns how to perform the required steps, their completion time will decrease, so there is a correlation between learning and completion time.
- The second experiment was conducted on a modified version of the simulator. The new version was based on the outcomes and on the feedback collected during the first experiment, which focused on (1) shortening the time of the fire extinguishing process at the first level, (2) reducing the navigation speed to slightly mitigate motion sickness, (3) introducing visual navigation guidance to help the trainees navigate through the virtual environments, as shown in Figure 11. The cue is a highlighted path with a text guiding the testers to their destination after the fire is discovered, and (4) providing visual feedback cues when the tasks are completed. In addition, a comprehensive task logger was added to the application to gain more insights into the behavior and workload of the trainees during the use of the simulator, as well as introducing emotion detection to complement the analysis of the workload. For this second experiment, 7 naval engineers (4 men, 3 women) from the SEAS 4.0 participated (3 of them tested the simulator for the first time). Like for the first experiment, the testers were introduced to the use of the controllers and on the fire drill protocols.

For both experiments, the testers were asked to provide their knowledge level and experience with fire drills, video games and VR, since such experience can influence the obtained results. Such data are shown in Tables 1 and 2.



Figure 11. Visual navigation cue with text with instructions.

Table 1. Background of the testers of the first experiment.

Tester No.	Gender	Previous Experience in Fire Drills	Previous Experience in VR	Previous Experience in Playing Video Games
1	Man	High	High	High
2	Man	High	Low	Low
3	Woman	Low	Low	High
4	Man	Medium	Low	Low
5	Woman	Low	Low	High
6	Man	Low	Low	High
7	Man	Low	Low	Low
8	Man	High	Low	High
9	Man	Low	Low	High
10	Man	High	Medium	High

Table 2. Background of the testers of the second experiment.

Tester No.	Gender	Previous Experience in Fire Drills	Previous Experience in VR	Previous Experience in Playing Video Games
1	Man	High	High	High
2	Woman	Medium	Low	High
3	Woman	High	Low	Low
4	Man	Medium	Medium	Medium
5	Man	Low	Low	Low
6	Woman	Medium	Medium	Low
7	Man	Medium	Medium	High

5.2. Performance Tests

5.2.1. Time Required to Complete Each Level

During the first experiment, as it was previously mentioned, the only collected data were the time to complete each level, which is depicted in Figure 12. As it can be observed, testers spent a lot of time on Level 1, which is essentially due to the fact that the implemented fire extinguishing process took a significant amount of time (52 s when performed perfectly), and since it was the first contact of the testers with the application, so they needed to adapt to the controls and to the virtual environment. A similar situation occurred in Level 3, where the testers needed to adapt to a new scenario (the engine room).

One conclusion that can be drawn from the experiments is that the testers with prior experience in video gaming were the fastest to complete the tasks and levels compared to the testers with low video gaming experience. For instance, an experienced gamer like Tester No. 1 required clearly less time than less experienced gamers. In fact, non-experienced gamers required on average roughly the double of the time of experienced gamers. Nevertheless, the impact of previous experience with fire drills or VR was less evident in the results.

Another observation concerning task completion is the fact that all testers followed the exact sequence of instructions provided by the fire drill, except for testers 4, 8, and 9. Specifically, testers 4 and 9 attempted to extinguish the fire in Level 2, while tester 8 evacuated before extinguishing the fire in Level 3. This suggested a need for a real-time task tracker (later implemented for the second experiment) to guide trainees when they do not accomplish a specific task. Nonetheless, overall, most testers improved their decision making skills in assessing fire severity and determining whether to extinguish the fire or to evacuate.

For the second experiment, in order to accelerate the training process, the time for extinguishing the fire was drastically reduced from 52 s to 7 s. In addition, the navigation speed was adjusted, and visual cues (a path) were added to enhance the guidance. With such modifications, the obtained times to finish each level were significantly cut down, as can be observed in Figure 13.

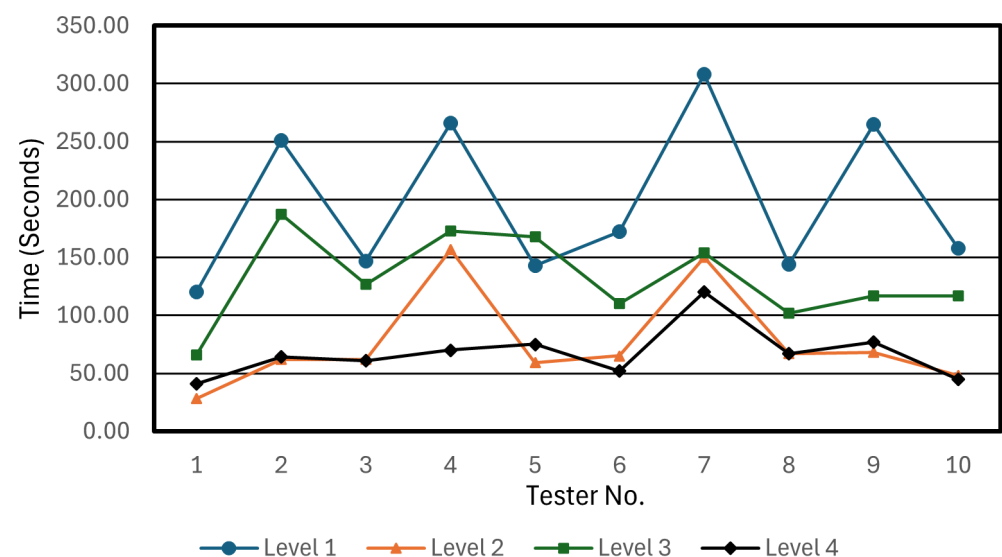


Figure 12. Time spent on each level for the first experiment.

It is important to note that Testers 2, 4, and 5 evaluated the simulator for the first time, so they required (especially testers 4 and 5) significantly more time than the other testers.

Moreover, it is worth pointing out that Testers 5 and 6 barely had any experience with video games and VR in comparison to Tester 2.

On the other hand, testers who participated in the first experiments performed better in the second experiment. This is due to the previous training, but also thanks to the additional help provided by the included additional visual guidance.

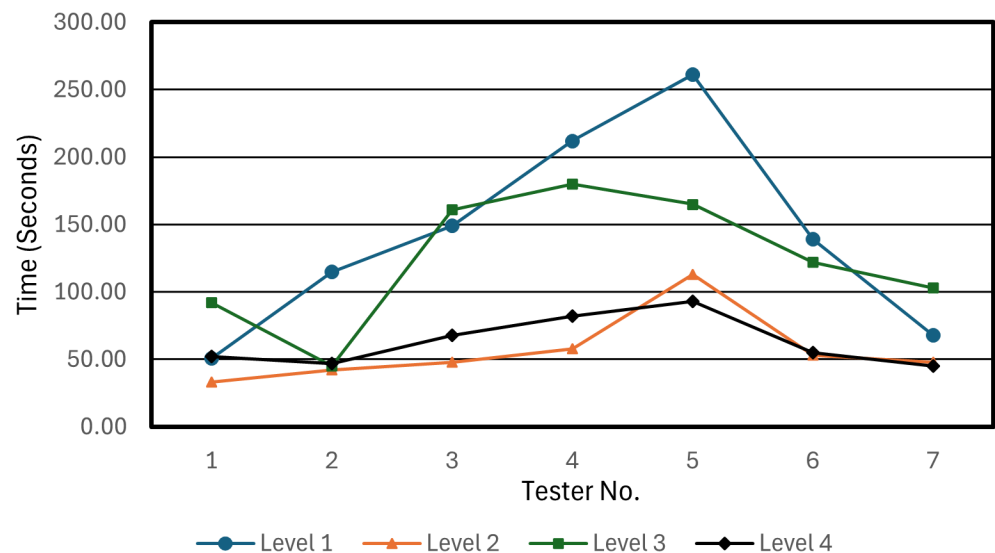


Figure 13. Time required to finish each level in the second experiment.

5.2.2. Number of Gazed Objects

For this comparison, the data from the second experiment were utilized to measure the amount of objects gazed by each tester. The gazes of Tester 1 were used as a reference due to their larger experience on onboard fire drill operation and VR. Moreover, the eye gaze data were pre-processed to eliminate the effect of eye blinking.

Figure 14 shows the number of gazes per tester. It can be observed that the testers stared at a lot of objects in levels 1 and 3, since such levels require exploring virtual environments that are new to the testers. Specifically, 19.38% of the gazes were aimed at the fire collider, 17.51% at the virtual phone collider, 14.46% at the fire extinguisher, and 10.7% at the fire alarm. This is normal because the previously mentioned objects are the main objects involved in the fire drill. However, it is worth noting that only 8.33% of the gazes were aimed at watching the signage (e.g., fire exits, muster area, emergency phone signs or the fire alarm sign) and 6.12% at the guiding text. This indicated a need for providing highlighted interactive elements on the signage and on the guiding text to attract the tester's gazes and to make the simulator more intuitive.

5.2.3. Sequence of Gazed Objects

In order to determine the similarity of the sequences of gazed objects, an approach used in bioinformatics for comparing genome sequences was used. Specifically, the Longest Common Subsequence (LCS) and the Sliding Window (SW) techniques were implemented. The LCS is the longest common subsequence between two sets of sequences. This method utilizes dynamic computing to construct a matrix with the length of the longest common subsequences [83]. This method finds the longest shared sequence of items between two different sequences, even if there are other items in between (this is illustrated in Figure 15, left). Thus, the similarity score with LCS is calculated by [84]:

$$\text{Similarity}_{\text{LCS}} = \frac{\text{LCS Length}}{\sqrt{\text{Ideal Length} \times \text{Compared Length}}} \quad (1)$$

where

- LCS Length is the length of the longest common subsequence between the sequence of gazed objects obtained by the reference tester (Tester 1 in these experiments) and the sequence of the tester to be compared.
- Ideal Length is the number of elements in the sequence of the reference tester.
- Compared Length is the number of elements of the sequence obtained by the tester to be compared.

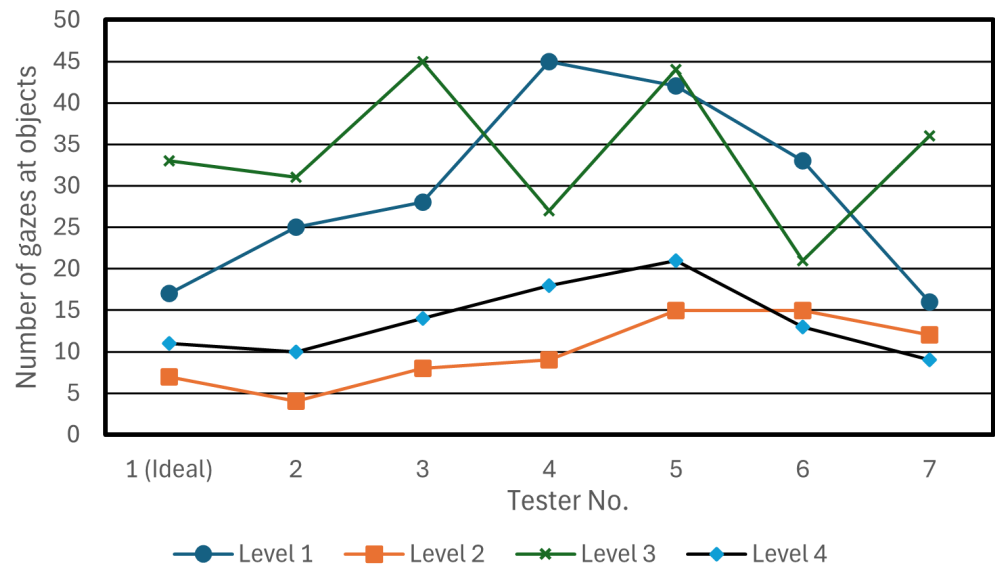


Figure 14. Number of the gazed objects for each tester and level.

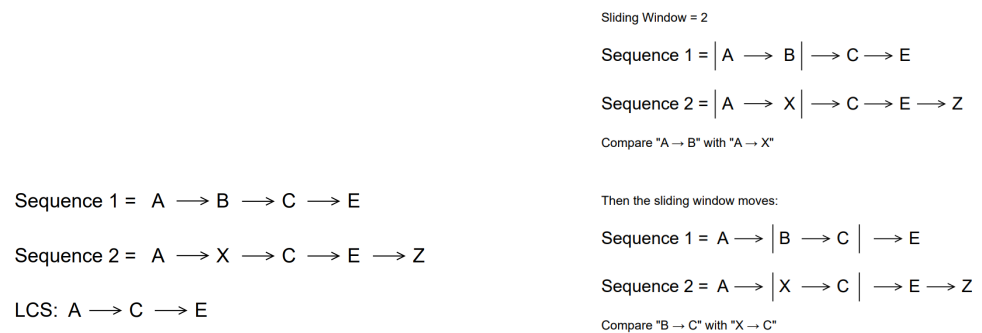


Figure 15. Example of the application of LCS to compare two sequences (left) and illustration of the Sliding Window technique (right).

On the other hand, SW offers a less strict approach to calculating the similarity score, where a window of a specified size is used to slide over the original sequence to find matches within the other sequence (this is illustrated in Figure 15, right). This approach is effective in identifying near matches rather than requiring perfect alignment [85]. Specifically, the similarity score for SW is computed as follows:

$$\text{Similarity}_{\text{SW}} = \frac{\text{Match Count}}{\sqrt{\text{Ideal Length} \times \text{Compared Length}}} \quad (2)$$

where:

- Match Count is the length of the longest common subsequence between the sequence of the reference tester and the sequence of the tester to be compared.
- Ideal Length is the number of elements in the sequence of the reference tester.
- Compared Length is the number of elements in the sequence of the tester to be compared.

Figure 16 shows the similarity score obtained when using LCS for the users that carried out the second experiments described in Section 5.1. In such a figure, it can be observed that the sequence similarity score of the levels tends to oscillate around 0.5: level 3 has an average similarity score of 0.552 and level 2 average score is 0.448, while levels 1 and 4 both average 0.506 and 0.508, respectively. Moreover, the average for each tester for all four levels fluctuates between 0.488 and 0.56.

It must be noted that the obtained results are conditioned by the definition of LCS, which is quite rigid (i.e., it is used to determine whether two sequences are exactly the same), when, in practice, during the experiments, the testers' gazes deviated from the ideal sequence due to the lack of indicators to follow a specific sequence of objects. Thus, this resulted in relatively low similarity scores when comparing with the ideal sequence.

Figure 17 shows the similarity scores obtained when using SW for a window size of 2 (such a size is necessary due to the non-colocalization of the gazed objects in the virtual environment). As can be observed in Figure 17, the average similarity score is lower than for LCS scores. This is essentially due to the small window size of 2: it requires strict close alignment precision, which is not flexible when gaze deviations occur.

Specifically, level 1 has the lowest average similarity score (0.292), while level 3 has the highest (0.472) (such a metric is 0.349 for level 2 and 0.425 for level 4). These results mean that levels 3 and 4 (the ones centered in the engine room) are more intuitive when navigating and interacting with objects.

Testers average a SW value between 0.285 and 0.506, which allows for estimating how far they are from learning the ideal sequence of actions to be taken in the simulated fire drill.

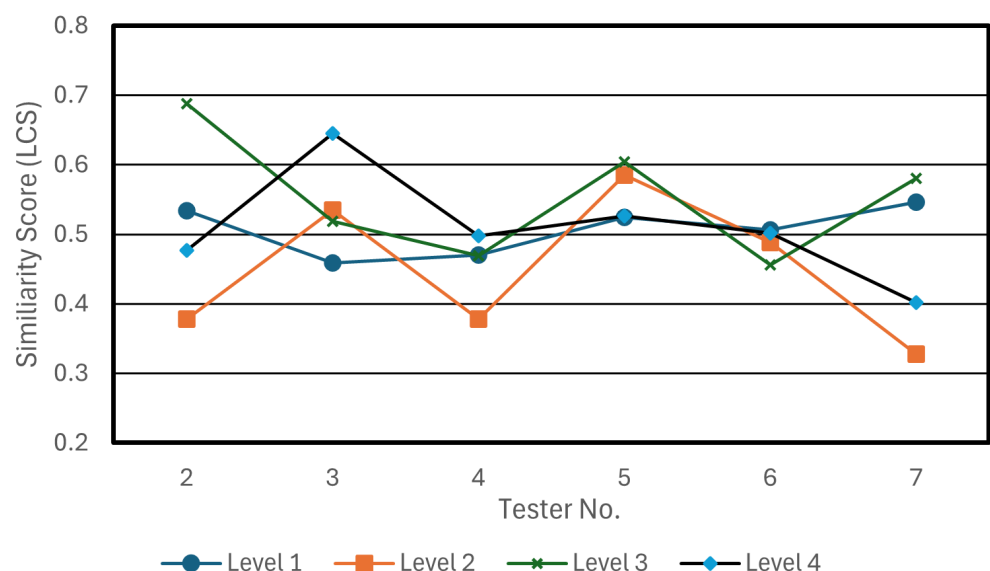


Figure 16. Similarity score for LCS.

Overall, after analyzing the LCS and SW values, it can be concluded that, in order to improve them, it is necessary to provide some sort of visual cues or guiding audio to help users during their training.

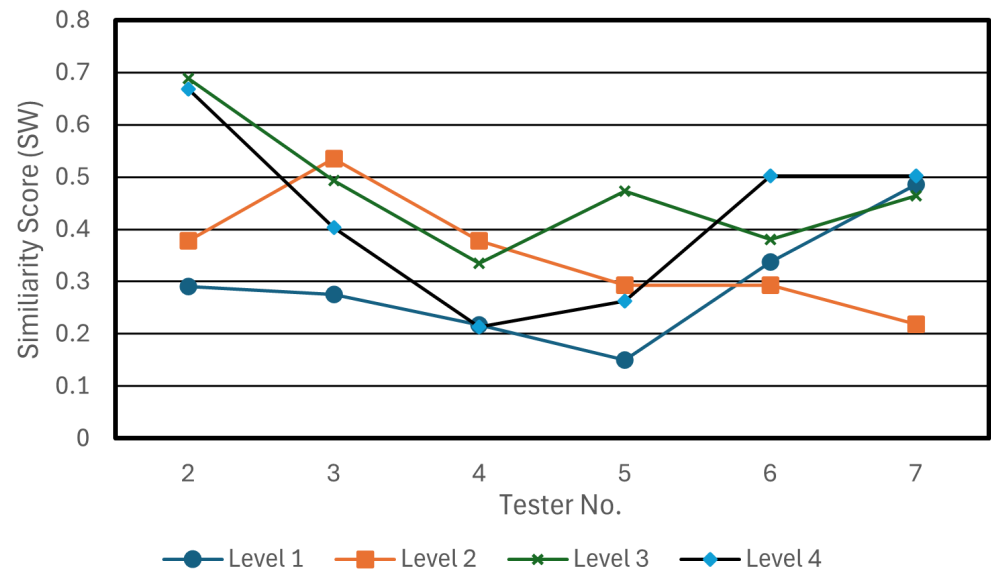


Figure 17. Similarity score for SW.

5.2.4. Performance with and Without Adding Visual Cues

To assess the performance improvement achieved after including additional visual navigation guidance (text and virtual arrows), the mean completion times and standard deviations for each level were compared. The results are shown in Table 3, where it can be observed that significant improvements were obtained. Specifically, on average, completion time was reduced between 14.18% and 32.72%, thus demonstrating that the included visual clues were clearly effective.

Moreover, Figure 18 shows the box plots that allow for illustrating the performance of the testers before (“old levels”) and after including additional visual guidance (“new levels”). As it can also be observed in Table 3, Level 2 shows the most significant decrease in both mean time and standard deviation.

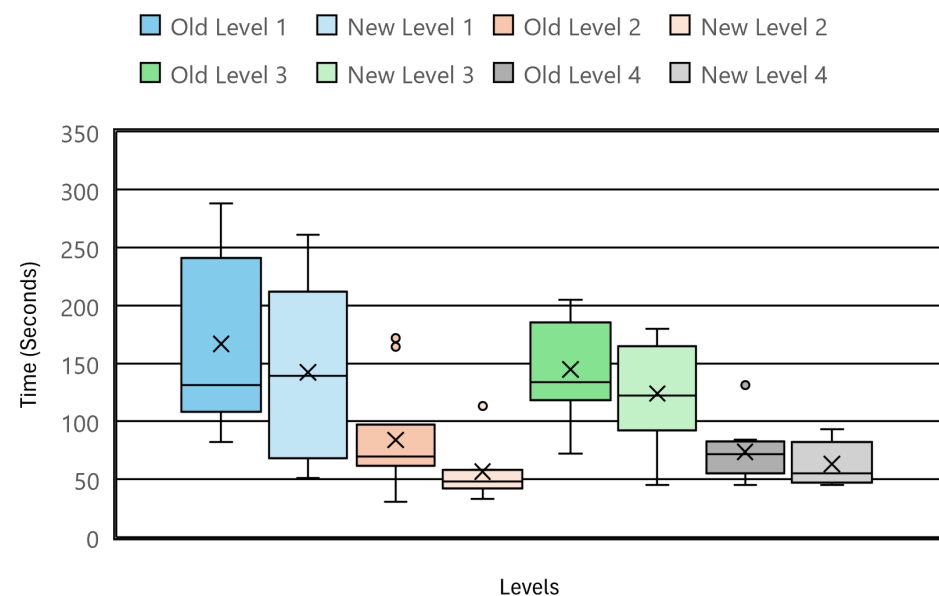


Figure 18. Box plot comparison of the performance of testers with and without the added visual cues.

Table 3. Completion times and standard deviation of old and new implementations.

Level	Old Implementation		New Implementation		% Improvement
	Mean (s)	Std. Dev (s)	Mean (s)	Std. Dev (s)	
1	166.88	73.84	142.14	74.84	14.82%
2	83.88	46.27	56.43	26.18	32.72%
3	144.65	41.13	124.00	48.12	14.31%
4	73.58	24.24	63.14	18.49	14.18%

5.3. Emotion Detection Performance

5.3.1. Experimental Setup

In order to evaluate the emotional performance of the testers, the FACS was utilized to interpret their raw expressions to detect valid emotions. The FACS is a system developed for classifying human facial expressions based on muscle movements. It was originally created by Paul Ekman and Wallace V. Friesen in 1978 [86] and is now widely used in psychology, neuroscience or artificial intelligence [87]. The framework actually identifies facial muscle movements that can be combined to infer emotions (e.g., happiness, sadness, anger, surprise). For instance, genuine smiles, which express happiness, are characterized by raising cheeks and pulling the corner of the lips, while wrinkling the nose and raising the upper lip are related to disgust.

The detected emotions were recorded in conjunction with the objects the testers were gazing at, which helped assess with accuracy the reactions of the testers in certain situations and when using specific objects. Thus, the detected facial expressions are used to interpret the level of “good” or “bad” emotions felt by each tester during the simulation.

It is important to note that, while the number of participants in this set of emotion detection experiments can be considered low, this is due to the fact that this article focus is on the technical feasibility of detecting emotions using XR smart glasses in a realistic immersive VR application. Therefore, providing an extensive analysis with a large group is beyond the scope of this article. Nonetheless, the emotion detection experiments described in the next sections provide valuable preliminary insights into user engagement and emotional response.

5.3.2. Emotion Detection Accuracy

To conduct this test, a list of facial expressions was mapped to each game object, and then assessed if the tester had made an expression that fit the list criteria. For instance, when a user looked at a fire, it was expected to detect fear. Thus, the more the tester reacted with an expression that fit the list criteria, the higher the emotion detection accuracy, which was measured by dividing the number of correctly detected expressions by the total number of detected expressions.

It is worth mentioning that, although the application was recording continuously the facial expressions of the testers, such expressions cannot always be related to an emotion as defined by the FACS. In such situations, the application recorded that the system detected “no emotion”. It is important to understand the difference between Figures 19 and 20, which show the different accuracy levels (i.e., the similarity of the detected emotion with respect to the one indicated in the list criteria in relation to a specific object) collected for all the participants in the second experiment described in Section 5.1. In practice, this means that the data depicted in Figure 19 was pre-processed to remove when “no emotion” was detected when staring at an object. Note that such a detection of “no emotion” can occur during the whole interaction with the object (which would count as incorrect emotion detection if that is different from what is indicated in the pre-established list of emotions

related to the object) or at some specific time instants during the interaction (for example, a tester can first express fear when using a fire extinguisher to put out a fire in the engine room, but then, as the fire is being put out, the tester may express some sort of relief that can be detected as “no emotion” by the Meta Quest Pro). In contrast, in Figure 20, such “no emotion” events are also accounted for, so the emotion detection accuracy shows significantly lower results with respect to Figure 19.

Considering the previously described experimental conditions, it can be observed in Figure 19 that Tester 1 shows the highest recorded emotion accuracy according to the established list criteria for all the levels when excluding the “no emotion” events. However, such a tester accuracy drops when including the “no emotion” events, which indicates that, although the tester reacted to the different situations, overall he did not express as many emotions as expected during the test. Regarding Tester 2, she expressed the expected emotions on levels 3 and 4, also achieving high accuracy when including the “no emotion” events.

With respect to the rest of the testers, although they expressed different emotions when using the application, they did not react as expected, so, overall, their accuracy levels in both Figures 19 and 20 are low. This is essentially related to how the predefined list of emotions was defined, which diverges from the actual reactions of the testers. For instance, when putting out a fire, it was expected to detect fear, but some testers actually showed anger and even happiness, as a sign of relief for performing the task correctly.

Finally, it was observed that, to obtain a good emotion detection accuracy, it was necessary to mount the Meta Quest Pro headset correctly, since incorrect mounting affects the capturing of facial expressions.

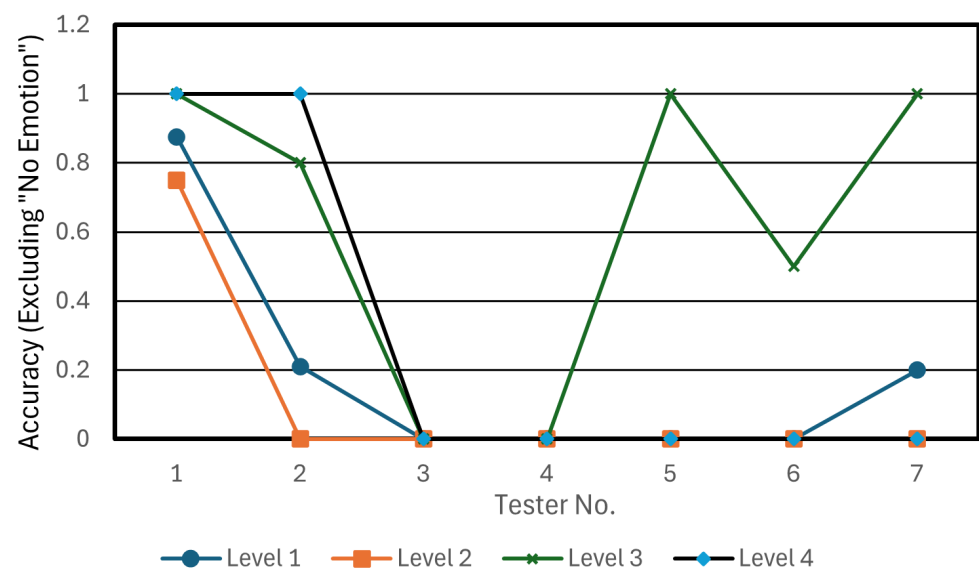


Figure 19. Emotion detection accuracy for each tester excluding “no emotion” events.

5.3.3. Emotions Detected for Each User and Level

The emotional responses captured during the performed fire drill simulations were categorized into three groups: good emotions (e.g., happiness, contempt), bad emotions (e.g., anger, fear, disgust), and no emotions expressed at all. These categories simplify the analysis, highlighting which scenarios are well-balanced, overly distressing or under-stimulating. Thus, the obtained results allow for the evaluation of the overall experience of each tester for each of the four levels of the simulation.

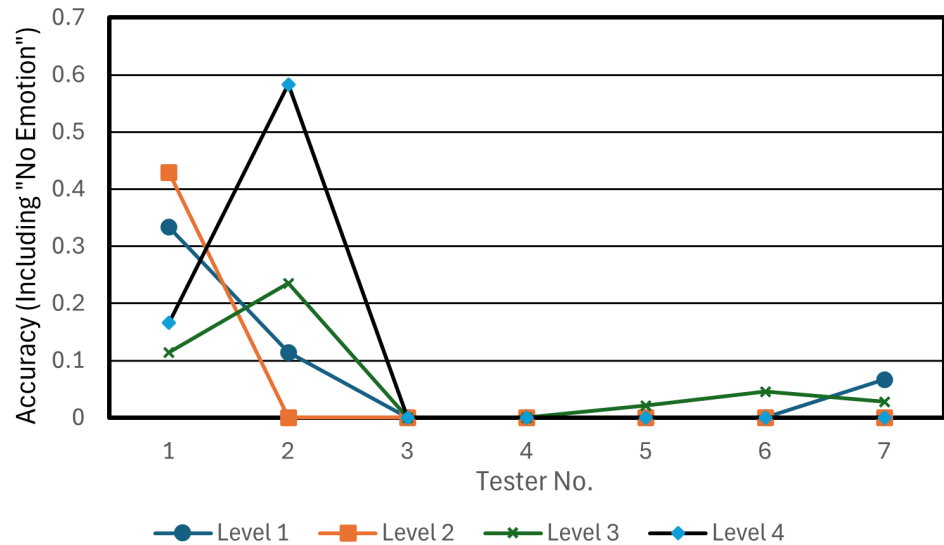


Figure 20. Emotion detection accuracy for each tester including “no emotion” events.

Figure 21 shows the percentage of detected good emotions for each user and level. It can be observed that positive emotions are, in general, very low for most of the levels. This was expected, since a ship fire drill involves a certain level of stress that cannot be directly linked to good feelings (except for situations where relief is expressed). In any case, there are some testers like 2 and 7 that, in Figure 21, show a significantly high level of good emotions (33.33%) in specific levels. Tester 4 also showed a relatively high percentage of good emotions (up to 14.58%). The rest of the testers had little to no positive emotions throughout the whole simulation.

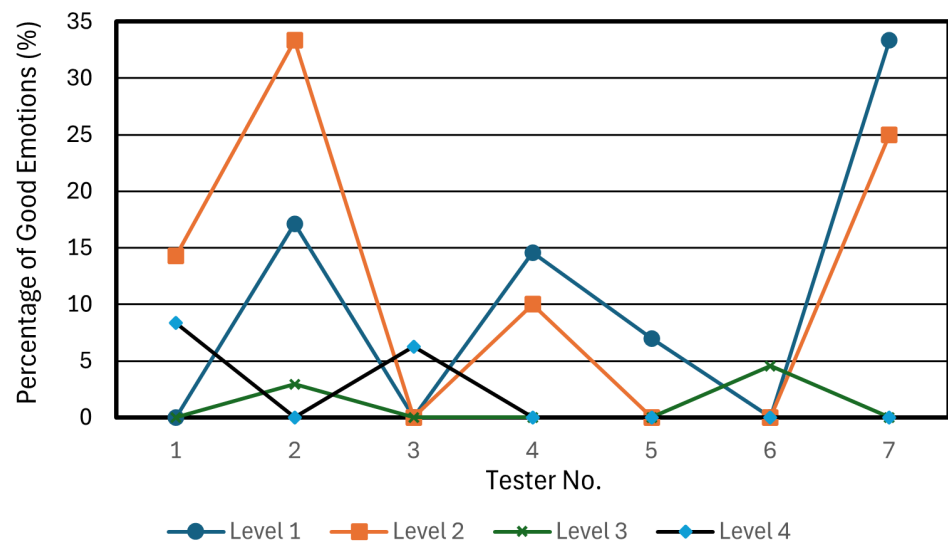


Figure 21. Percentage of good emotions experienced by each tester through each level.

On the other hand, negative feelings were more prominent, as demonstrated in Figure 22. For instance, Tester 2 showed the highest level of negative emotional states, with 66.67% at level 2 and 58.33% at level 4. Tester 1 was close behind, with 42.86% at level 2 and 38.10% at level 1. Testers 3 and 7 had only average levels of negative emotions, while testers 4, 5 and 6 displayed little to no negative emotional responses.

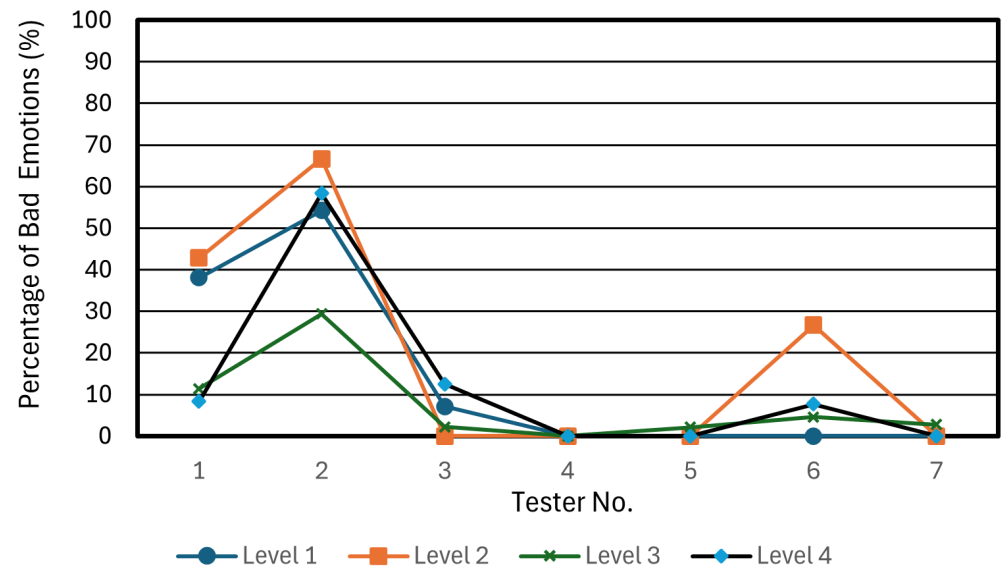


Figure 22. Percentage of bad emotions experienced by each tester through each level.

Finally, Figure 23 shows that a significant proportion of the testers expressed high intensities of no emotions across all rating levels. For instance, Tester 3 exhibited a lack of emotion exceeding 90% across all levels, with the exception of level 4, where the percentage fell to 81.25%. Likewise, Testers 5 and 6 consistently recorded a 100% absence of emotion in multiple levels. In contrast, Testers 1 and 2 demonstrated diminished levels of emotional detachment as the simulation advanced, with Tester 1 decreasing from 61.90% in level 1 to 42.86% in level 2, suggesting a rise in emotional involvement.

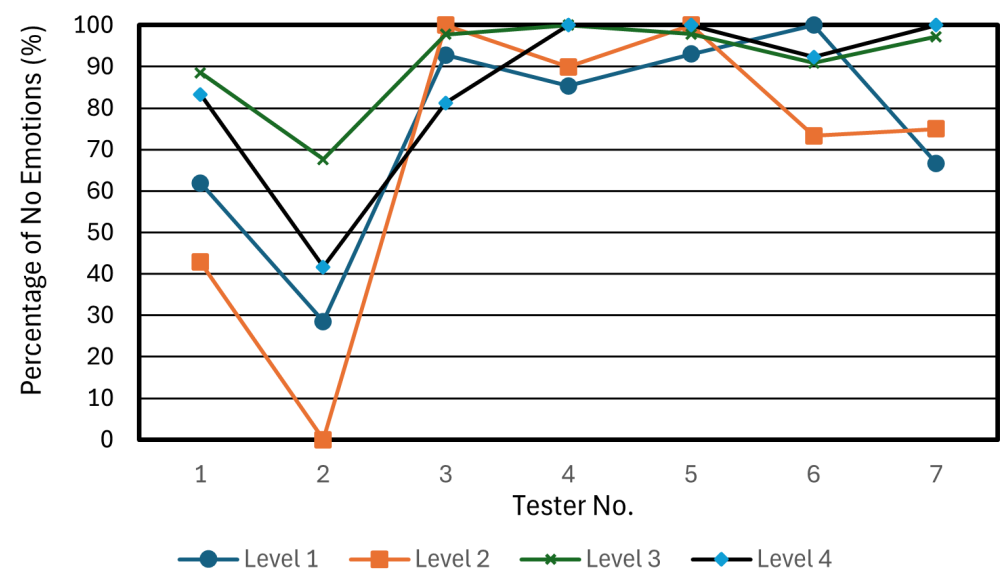


Figure 23. Percentage of “No Emotion” captured for each tester through each level.

Overall, the obtained results suggest that there was a clear division between the testers who felt emotional during the simulation and those who did not. Such a lack of emotional involvement can be associated with multiple factors related to each individual (e.g., personality traits, empathy, risk perception, prior exposure to real-life emergency situations), as well as to the fact that the testers considered the task more like a video-game rather than an actual fire drill, which affected their sense of presence (their awareness of

being inside a virtual environment) [88]. In addition, since some of the testers retook the tests after the first experiment, they were aware of the fire drill scenario, which may have derived into a degree of emotional desensitization, where familiarity with the scenario lessened the perceived urgency or seriousness of the fire drill.

Moreover, it is worth indicating that this article is focused more on describing the design, development and demonstration of the viability of detecting emotions in a practical application, so the size of the test group and their profile (e.g., naval engineers) condition the potential results. Thus, although it is out of the scope of this article, future research should be aimed at evaluating larger and more diverse testing groups.

6. Key Findings and Answers to the Proposed RQs

The path to the Metaverse still needs to be paved to provide truly immersive experiences that consider humans as their center, both in terms of user experience (i.e., interactions) and real response (i.e., reactions). After carrying out the development and evaluation of the ship fire drill application, the following key findings can be indicated to future researchers and Metaverse developers in relation to the analyzed research questions:

- RQ1: How can emotion awareness be integrated into XR and Metaverse applications to improve user interaction, particularly in training simulations?
- As it was described throughout this article, emotional awareness can be integrated into training simulations to potentially enhance user interaction. In the specific use case presented, facial expressions were used to interpret users' emotional states. In mission-critical contexts like emergency drills, this approach increases realism, enables personalized feedback, and supports the assessment of both technical performance and emotional resilience. Nonetheless, a number of challenges still remain, as it is later discussed in Section 7.
- RQ2: What is the current gap in traditional XR and Metaverse applications regarding the role of emotions in user experience?
- The current gap in traditional XR and Metaverse applications lies in their limited ability to recognize, interpret and respond to users' emotional states in real time.
 - Although technologies like VR have progressed significantly in recent years and have become more affordable, they still need to be enhanced in terms of graphics to provide more realism in certain situations. Moreover, such technologies need to be enhanced to overcome a problem that is more concerning: during the two performed experiments, five testers experienced slight to severe motion sickness. This indicates that completely immersive technologies may not be adequate for certain people or they require some level of adaptation to use them.
 - Emotional response is not easy to predict. An example of this fact is illustrated in Table 4, where there is a notable difference between the expected emotional responses and the ones detected when the testers were tasked with "Locating the fire". Of the seven testers, only Tester 1 was surprised during all four levels of the test. Tester 2 went from an indifferent attitude to a fearful one in the engine room setting. The rest of the testers showed indifference throughout the simulation, which is opposite to the expected fear or surprise. This indicates that the simulator did not induce a strong emotional response in the testers, but if that is a requirement for other applications, additional immersive artifacts should be included.

- The analysis of aggregated emotional responses, as summarized in Table 5, reveals a significant divergence between expected and observed emotions across all simulated tasks. With an overall match rate of only 18.8%, and a predominant observed emotional state of “indifferent” (accounting for 81.2% of observations where specific emotions like happiness, fear, or surprise were anticipated), the findings suggest that the current iteration of the simulator did not consistently elicit the targeted emotional engagement necessary for these critical training scenarios.
- In relation to the previous points, it is worth noting that every tester showed different responses in the same fire drill scenario, even when they were experiencing a similar situation. The results show that there were testers who showed no specific emotions, but it must be noted that such a response is conditioned by the scenario, the personality and background of each tester, the correct mounting of the headset and by how FACS defines emotion detection. These factors allow for concluding that, although headsets like Meta Quest Pro enable detecting facial expressions, it is necessary to calibrate the system for every person before its actual training to detect accurately emotions. Moreover, external hardware (e.g., IoT sensors) can also be a good complement to facial expression detection.

RQ3: What architectural components are necessary for developing an emotion-aware Metaverse application, specifically for real-time emotion detection?

- First, it is necessary to provide a sensing layer to collect data such as facial expressions, voice cues or physiological signals through devices like VR headsets, microphones, wearables, haptic suits, and/or IoT devices. These data are then processed by an emotion recognition module that can make use of rule-based or AI models to analyze and classify emotional states/behavior based on established frameworks to infer the trainee’s emotional and cognitive state continuously. The interpreted emotions feed into a context-aware adaptation engine, which can dynamically adjust the virtual environment by modifying parameters (e.g., difficulty, virtual object responses or context-awareness settings) to create a personalized and responsive user experiences. All these components can be integrated through a middleware layer that ensures real-time data flow, low latency and compliance with privacy and security standards. Finally, the used XR interface would be responsible for rendering the immersive environment and reflects these adaptations.

RQ4: How can eye-tracking and facial expression analysis be incorporated into XR headsets (e.g., Meta Quest Pro) to facilitate emotion detection in a Metaverse training environment?

- As it was described in this article, eye-tracking, and facial expression analysis were successfully integrated into the Meta Quest Pro smart glasses by leveraging its built-in sensors and cameras, designed to capture detailed facial movements and gaze patterns. Devices of this kind typically feature infrared eye-tracking cameras that monitor pupil position, blink rate and gaze direction, along with inward-facing cameras that detect subtle facial muscle movements.

RQ5: Can trainees’ performance be improved in a VR fire drill simulator by adding design elements and UI refinements?

- Yes, thoughtful design elements and UI refinements (e.g., real-time task tracker, clearer visual cues) in the VR fire drill simulator significantly im-

proved trainee performance, particularly in terms of task completion time. The results showed that these design improvements reduced task completion times by between 14.18% and 32.72%.

RQ6: Can the integration of real-time emotion detection in a VR fire drill simulator provide valuable insights into trainees' responses under stress?

- It was demonstrated that it is currently possible to create applications for training future Meta-Operators in critical tasks like disaster relief scenarios where fires occur.
- Although the performed tests allowed for illustrating that a future emotion-aware Metaverse can be possible, the limited sample size impedes the extraction of deeper conclusions. For instance, such a reduced size influences the conclusions drawn from the second experiment, since the testers who retook this simulation seemed less emotionally involved than those taking it for the first time. This could be interpreted as a bit of emotional desensitization, wherein continued access to the simulation may have diminished a perception of urgency or seriousness associated with a fire scenario. It is important to note that other factors, like cultural and social norms, also affect the emotional responses obtained by a test group [89].
- To obtain better performance results, future application developers should consider that the higher the realism, the better the engagement of the users and the lesser the approach to the simulator as a form of entertainment. The prospects for realism can be enhanced by the development of multisensory feedback and personalization. Developers should also introduce variability to prevent desensitization and prioritize emotional realism through dynamic AI-driven interactions. By focusing on these areas, future virtual environments can foster stronger emotional connections and more impactful experiences.

Table 4. Testers' emotional responses when locating and extinguishing the fire across different levels.

Tester No.	Level 1	Level 2	Level 3	Level 4	Expected Emotion
1	Surprise	Surprise	Surprise	Surprise	Surprise/Fear
2	Indifferent	Indifferent	Fear	Fear	Surprise/Fear
3	Indifferent	Indifferent	Indifferent	Indifferent	Surprise/Fear
4	Indifferent	Indifferent	Indifferent	Indifferent	Surprise/Fear
5	Indifferent	Indifferent	Indifferent	Indifferent	Surprise/Fear
6	Indifferent	Indifferent	Indifferent	Indifferent	Surprise/Fear
7	Indifferent	Indifferent	Indifferent	Indifferent	Surprise/Fear

Table 5. Alignment of observed emotions with expected emotions per task.

Task	Expressed Emotion	# Observations	# Matches	Match (%)	# Mismatch	Mismatch (%)	Predominant Emotion
Arriving at the muster area	Happiness/content	28	5	17.9%	23	82.1%	Indifferent
Calling the ship master	Fear/surprise	28	4	14.3%	24	85.7%	Indifferent
Locating the fire	Fear/surprise	28	7	25.0%	21	75.0%	Indifferent
Triggering the fire alarm	Fear/surprise	28	5	17.9%	23	82.1%	Indifferent
Overall	Varied	112	21	18.8%	91	81.2%	Indifferent

7. Limitations of the Study and Future Work

The work presented in this article evaluated the technical feasibility of making use of XR smart glasses to detect emotions in realistic Metaverse applications, but the potential conclusions drawn from the provided analyses are impacted by several factors:

- **Experimental group size:** this article focused on the mentioned evaluation of the technical feasibility of detecting emotions rather than providing an extensive analysis with a large experimental group. Therefore, the presented analyses should be considered just as an example of the responses from a representative but reduced group. Thus, future research should be aimed at evaluating larger groups with enough diversity to draw accurate conclusions. In addition, incorporating a qualitative assessment such as a 5-point Likert-scale questionnaire, would be advisable to provide deeper psychological insight and further explain the results.
- **Experimental group experience and profile:** the selected experimental group had a very specific background on naval engineering and an overall good knowledge on the use of recent technologies. However, this is not the case of many ship crews, so future research should also include people with diverse backgrounds and previous experience.
- **Effectiveness of the developed application:** although special care was put into the development of the dire drill application, the way certain visual and auditory effects are implemented impact the trainees in a different way depending on their individual emotional response. Therefore, future work should consider such diversity, and ideally, adapt the implementation to make it more effective.
- **Used hardware and software:** the selected XR smart glasses can be considered state-of-the-art hardware, but they can still be considered as a sort of commercial prototype in terms of emotion detection. Thus, their use and the provided APIs conditioned the performed experiments and the obtained results. Future researchers should monitor the Metaverse-ready XR market and keep an eye on the latest hardware and software, which has evolved really quickly in recent years.
- **Scenario of the proposed Metaverse application:** Real-life shipboard fire emergencies require a calm, controlled and well-trained response, often minimizing the influence of emotional reactions. It is important to recognize the inherent limitations of simulations in fully replicating the intensity, urgency, and unpredictability of such real-world emergencies. The goal of incorporating emotional tracking in this context is not to mimic the emotional state of seasoned professionals during actual emergencies, but rather to gain valuable insights into how less experienced individuals such as trainees emotionally respond within realistic Metaverse applications. A clear understanding of such reactions can help identify critical stress points, design more effective training scenarios and applications and ultimately support more effective training strategies.

Regarding other aspects of future research, they can be focused on enhancing a fire drill simulator realism, user experience and emotional impact. Key development areas include

- **Enhanced emotional realism:** multi-sensory feedback, such as dynamic, spatially-aware audio and congruent haptic sensations (e.g., localized heat, vibrotactile feedback), can be integrated to create stronger and more predictable emotional responses.
- **Improved visuals and interaction:** to address graphical limitations and motion sickness, visual and environmental realism can be improved by using advanced rendering (e.g., Ray Tracing), detailed interactions or Computational Fluid Dynamics (CFD) fire models. Locomotion and interaction mechanics can be refined to minimize adverse effects and enhance user engagement.

- **Increased scenario variability and engagement:** to prevent emotional desensitization and foster stronger emotional connections, significant scenario variability can be introduced. This includes multiplayer scenarios for realistic team dynamics [44] and dynamic AI-driven interactions, which can potentially leverage real fire accident data to sustain emotional realism and engagement.
- **Robust emotion detection:** to counter variability in emotional expression, advanced systems for emotion detection can be developed. Machine learning models (e.g., temporal classifiers like Long-Short Term Memory (LSTM) or dimensional emotion regressors) may be employed as an alternative to traditional rule-based FACS interpretation. These systems are able to include individual calibration capabilities. In addition, complementary external hardware, such as Internet of Things (IoT)-based physiological sensors, can be integrated for a more holistic understanding of trainee emotional states.

8. Conclusions

This article introduced an XR fire drill simulator designed for maritime safety training that is able to monitor the interactions of the users, what they stare at and to detect their emotions. By integrating emotion detection technologies such as eye tracking and facial expression analysis, the system aimed to enhance realism, evaluate trainees' stress responses and improve decision making in high-pressure fire emergencies.

After providing a detailed review of the state of the art and background work, the article described the main components of the developed emotion-aware XR application. The performed experiments allowed for measuring the amount of time spent in carrying out certain tasks, concluding that trainees with prior VR or gaming experience navigated the environment more efficiently. Moreover, the introduction of task-tracking visuals and navigation guidance significantly improved user performance, reducing the task completion times between 14.18% and 32.72%, thus enhancing procedural adherence. Emotional response varied widely among participants: while some exhibited engagement, others remained indifferent, likely treating the experience as a game rather than a critical safety drill. In addition, the carried out emotion-based performance analysis indicates that fear and surprise were not consistently elicited as expected, suggesting a need for more immersive elements.

In conclusion, this article represents a first step toward the development of emotion-aware Metaverse applications, offering initial insights through the description of a fire drill simulator that could guide future researchers and developers.

Author Contributions: Conceptualization, T.M.F.-C.; methodology, P.F.-L. and T.M.F.-C.; investigation, M.H.H.-A., D.R.-L., P.F.-L. and T.M.F.-C.; writing—original draft preparation, M.H.H.-A., D.R.-L., P.F.-L. and T.M.F.-C.; writing—review and editing, M.H.H.-A., D.R.-L., P.F.-L. and T.M.F.-C.; supervision, P.F.-L. and T.M.F.-C.; project administration, T.M.F.-C.; funding acquisition, P.F.-L. and T.M.F.-C. All authors have read and agreed to the published version of the manuscript.

Funding: This work has been funded by grant TED2021-129433A-C22 (HELENE) funded by MCIN/AEI/10.13039/501100011033 and the European Union NextGenerationEU/PRTR. In addition, this work has been supported by Centro Mixto de Investigación UDC-NAVANTIA (IN853C 2022/01), funded by GAIN (Xunta de Galicia) and ERDF Galicia 2021–2027.

Institutional Review Board Statement: This study involves only fully anonymized data and does not include any collection of personal or sensitive data. The study did not require ethical approval.

Informed Consent Statement: Verbal informed consent was obtained from the participants. The rationale for utilizing verbal consent is that the study involved minimal risk, did not include the

collection of sensitive personal data, and was conducted in an academic context. Verbal consent ensured accessibility and voluntary participation while maintaining ethical standards.

Data Availability Statement: Data is contained within the article.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Farooq, M.S.; Ishaq, K.; Shoaib, M.; Khelifi, A.; Atal, Z. The Potential of Metaverse Fundamentals, Technologies, and Applications: A Systematic Literature Review. *IEEE Access* **2023**, *11*, 138472–138487. [\[CrossRef\]](#)
2. Al-Ghaili, A.M.; Kasim, H.; Al-Hada, N.M.; Hassan, Z.B.; Othman, M.; Tharik, J.H. A Review of Metaverse's Definitions, Architecture, Applications, Challenges, Issues, Solutions, and Future Trends. *IEEE Access* **2022**, *10*, 125835–125866. [\[CrossRef\]](#)
3. Poncela, A.V.; Fraga-Lamas, P.; Fernández-Caramés, T.M. On-Device Automatic Speech Recognition for Low-Resource Languages in Mixed Reality Industrial Metaverse Applications: Practical Guidelines and Evaluation of a Shipbuilding Application in Galician. *IEEE Access* **2025**, *13*, 77017–77038. [\[CrossRef\]](#)
4. Li, K.; Lau, B.P.L.; Yuan, X.; Ni, W.; Guizani, M.; Yuen, C. Toward Ubiquitous Semantic Metaverse: Challenges, Approaches, and Opportunities. *IEEE Internet Things J.* **2023**, *10*, 21855–21872. [\[CrossRef\]](#)
5. Fraga-Lamas, P.; Lopes, S.I.; Fernández-Caramés, T.M. Towards a Blockchain and Opportunistic Edge Driven Metaverse of Everything. *arXiv* **2024**, arXiv:2410.20594.
6. Gu, J.; Wang, J.; Guo, X.; Liu, G.; Qin, S.; Bi, Z. A Metaverse-Based Teaching Building Evacuation Training System With Deep Reinforcement Learning. *IEEE Trans. Syst. Man Cybern. Syst.* **2023**, *53*, 2209–2219. [\[CrossRef\]](#)
7. Hamed-Ahmed, M.H.; Fraga-Lamas, P.; Fernández-Caramés, T.M. Towards the Industrial Metaverse: A Game-Based VR Application for Fire Drill and Evacuation Training for Ships and Shipbuilding. In Proceedings of the 29th International ACM Conference on 3D Web Technology, Guimarães, Portugal, 25–27 September 2024; pp. 1–6.
8. Wu, J.; Jin, Y.; Fu, J. Effectiveness evaluation on fire drills for emergency and PSC inspections on board. *TransNav Int. J. Mar. Navig. Saf. Sea Transp.* **2014**, *8*, 229–236. [\[CrossRef\]](#)
9. Stepien, J.; Pilarska, M. Selected operational limitations in the operation of passenger and cargo ships under SOLAS Convention (1974). *Zesz. Nauk. Akad. Morskiej Szczecinie* **2021**, *65*, 21–26.
10. Stephenson, N. *Snow Crash*; Bantam Books: New York, NY, USA, 1992; 440p, ISBN 10.0553351923.
11. Maier, M.; Hosseini, N.; Soltanshahi, M. INTERBEING: On the Symbiosis between INTERNet and Human BEING. *IEEE Consum. Electron. Mag.* **2023**, *13*, 98–106. [\[CrossRef\]](#)
12. Ritterbusch, G.D.; Teichmann, M.R. Defining the Metaverse: A Systematic Literature Review. *IEEE Access* **2023**, *11*, 12368–12377. [\[CrossRef\]](#)
13. Kontogianni, E.; Anthopoulos, L. Towards a Standardized Metaverse Definition: Empirical Evidence From the ITU Metaverse Focus Group. *IEEE Eng. Manag. Rev.* **2025**, *53*, 110–127. [\[CrossRef\]](#)
14. Mosco, V. Into the Metaverse: Technical Challenges, Social Problems, Utopian Visions, and Policy Principles. *Javnost-Public* **2023**, *30*, 161–173. [\[CrossRef\]](#)
15. Chai, Y.; Qian, J.; Younas, M. Metaverse: Concept, Key Technologies, and Vision. *Int. J. Crowd Sci.* **2023**, *7*, 149–157. [\[CrossRef\]](#)
16. Deloitte China, 'Metaverse Report: Future Is Here; Global XR Industry Insight', Mar. 2022. Available online: <https://www2.deloitte.com/cn/en/pages/technology-media-and-telecommunications/articles/metaverse-whitepaper.html> (accessed on 1 March 2025).
17. Blanco-Novoa, Ó.; Fraga-Lamas, P.; Vilar-Montesinos, M.A.; Fernández-Caramés, T.M. Towards the Internet of Augmented Things: An Open-source Framework to Interconnect IoT Devices and Augmented Reality Systems. *Proceedings* **2019**, *15*, 30.
18. Elias, D.; Ziegenbein, D.; Mundhenk, P.; Hamann, A.; Rowe, A. The Cyber-Physical Metaverse-Where Digital Twins and Humans Come Together. In Proceedings of the Design, Automation & Test in Europe Conference & Exhibition (DATE), Antwerp, Belgium, 17–19 April 2023.
19. Negri, E.; Abdel-Aty, T.-A. Clarifying concepts of Metaverse, Digital Twin, Digital Thread and AAS for CPS-based production systems. *IFAC-PapersOnLine* **2023**, *56*, 6351–6357. [\[CrossRef\]](#)
20. Jamshidi, M.; Serej, A.D.; Jamshidi, A.; Moztarzadeh, O. The Meta-Metaverse: Ideation and Future Directions. *Future Internet* **2023**, *15*, 252. [\[CrossRef\]](#)
21. Buchholz, F.; Oppermann, L.; Prinz, W. There is more than one metaverse. *I-com* **2022**, *21*, 313–324. [\[CrossRef\]](#)
22. IEEE CTS/MSC-Metaverse Standards Committee. P2048-Standard for Metaverse: Terminology, Definitions, and Taxonomy. Available online: <https://standards.ieee.org/ieee/2048/11169/> (accessed on 24 September 2024).
23. Fernández-Caramés, T.M.; Fraga-Lamas, P. Forging the Industrial Metaverse for Industry 5.0: Where Extended Reality, IIoT, Opportunistic Edge Computing, and Digital Twins Meet. *IEEE Access* **2024**, *12*, 95778–95819. [\[CrossRef\]](#)

24. International Maritime Organization (IMO). International Convention for the Safety of Life at Sea (SOLAS). 1974. Available online: [https://www.imo.org/en/About/Conventions/Pages/International-Convention-for-the-Safety-of-Life-at-Sea-\(SOLAS\)-1974.aspx](https://www.imo.org/en/About/Conventions/Pages/International-Convention-for-the-Safety-of-Life-at-Sea-(SOLAS)-1974.aspx) (accessed on 2 March 2025).
25. IMO. SOLAS, Consolidated Edition, 2020: Consolidated Text of the International Convention for the Safety of Life at Sea, 1974, and Its Protocol of 1988: Articles, Annexes and Certificates; International Maritime Organization: London, UK, 2020.
26. Sim, H.; Ha, W.-J.; Park, Y.-S. A basic study on standardization of fire-fighting drill scenarios on board. *J. Int. Marit. Safety, Environ. Aff. Shipp.* **2019**, *3*, 28–35. [\[CrossRef\]](#)
27. Cárdenas-Robledo, L.A.; Hernández-Uribe, O.; Reta, C.; Cantoral-Ceballos, J.A. Extended reality applications in industry 4.0.—A systematic literature review. *Telemat. Inform.* **2022**, *73*, 101863. [\[CrossRef\]](#)
28. Vukelić, G.; Vizentin, G.; Frančić, V. Prospects for use of extended reality technology for ship passenger evacuation simulation. *Pomorstvo* **2021**, *35*, 49–56. [\[CrossRef\]](#)
29. Wang, J.; Zhu, M.; Fan, X.; Yin, X.; Zhou, Z. Multi-channel augmented reality interactive framework design for ship outfitting guidance. *IFAC-PapersOnLine* **2020**, *53*, 189–196. [\[CrossRef\]](#)
30. Miličević, M.; Vidović, M.S.; Bonković, M.; Kuzmanić, A.; Ljubenkov, B.; Mimica, R.; Željковиć, I. Augmented reality for an efficient inspection of outfitting process in shipbuilding. In Proceedings of the 2021 International Conference on Software, Telecommunications and Computer Networks (SoftCOM), Split Hvar, Croatia, 23–25 September 2021; pp. 1–4.
31. Vidal-Balea, A.; Blanco-Novoa, O.; Fraga-Lamas, P.; Vilar-Montesinos, M.; Fernández-Caramés, T.M. A collaborative augmented reality application for training and assistance during shipbuilding assembly processes. *Proceedings* **2020**, *54*, 4. [\[CrossRef\]](#)
32. Morikawa, K.; Ando, T. Reduction of piping management person-hours through use of AR technology at shipbuilding sites. *Fujitsu Sci. Tech. J.* **2019**, *55*, 20–26.
33. Vidal-Balea, A.; Blanco-Novoa, O.; Fraga-Lamas, P.; Vilar-Montesinos, M.; Fernández-Caramés, T.M. A collaborative industrial augmented reality digital twin: Developing the future of shipyard 4.0. In *International Summit Smart City 360*; Springer International Publishing: Cham, Switzerland, 2021; pp. 104–120.
34. Mystakidis, S.; Besharat, J.; Papantzikos, G.; Christopoulos, A.; Stylios, C.; Agorgianitis, S.; Tselentis, D. Design, Development, and Evaluation of a Virtual Reality Serious Game for School Fire Preparedness Training. *Educ. Sci.* **2022**, *12*, 281. [\[CrossRef\]](#)
35. Morélot, S.; Garrigou, A.; Dedieu, J.; N'kaoua, B. Virtual Reality for Fire Safety Training: Influence of Immersion and Sense of Presence on Conceptual and Procedural Acquisition. *Comput. Educ.* **2021**, *166*, 104145. [\[CrossRef\]](#)
36. Haavardtun, P.; Mallam, S.; Sharma, A.; Nazir, S. Using Virtual Worlds as an Integrated Part of Virtual Distance Learning (VDL) Networks in Simulator-Based Education. In *Advances in Human Factors in Training, Education, and Learning Sciences*; Nazir, S., Ahram, T., Karwowski, W., Eds.; Advances in Intelligent Systems and Computing; Springer: Cham, Switzerland, 2020; Volume 1211, pp. 302–309. [\[CrossRef\]](#)
37. Wahidi, S.I.; Pribadi, T.W.; Rajasa, W.S.; Arif, M.S. Virtual reality-based application for safety training at shipyards. *IOP Conf. Ser. Earth Environ. Sci.* **2022**, *972*, 012025. [\[CrossRef\]](#)
38. Lee, G.A.; Yang, U.; Son, W.; Kim, Y.; Jo, D.; Kim, K.; Choi, J.S. Virtual reality content-based training for spray painting tasks in the shipbuilding industry. *ETRI J.* **2010**, *32*, 695–703. [\[CrossRef\]](#)
39. Wu, H.; Yang, J.; Chen, C.; Wan, Y.; Zhu, X. Research of Virtual Ship Fire-fighting Training System Based on Virtual Reality Technique. *IOP Conf. Ser. Mater. Sci. Eng.* **2019**, *677*, 042100. [\[CrossRef\]](#)
40. Markopoulos, E.; Lauronen, J.; Luimula, M.; Lehto, P.; Laukkanen, S. Maritime safety education with VR technology (MarSEVR). In Proceedings of the 2019 10th IEEE InternMystakidisational Conference on Cognitive Infocommunications (CogInfoCom), Naples, Italy, 23–25 October 2019; pp. 283–288.
41. Luimula, M.; Haavisto, T.; Pham, D.; Markopoulos, P.; Aho, J.; Markopoulos, E.; Saarinen, J. The use of metaverse in maritime sector—A combination of social communication, hands on experiencing and digital twins. *Creat. Innov. Entrep.* **2022**, *31*, 115–123.
42. Makransky, G.; Klingenberg, S. Virtual reality enhances safety training in the maritime industry: An organizational training experiment with a non-WEIRD sample. *J. Comput. Assist. Learn.* **2022**, *38*, 1127–1140. [\[CrossRef\]](#)
43. Shaoyang, Q.; Hongxiang, R. Ship life-saving training system based on virtual reality technology. In Proceedings of the 2018 IEEE 4th International Conference on Control Science and Systems Engineering (ICCSSE), Wuhan, China, 21–23 August 2018; pp. 559–563.
44. Ginnis, A.I.; Kostas, K.V.; Politis, C.G.; Kaklis, P.D. VELOS: A VR platform for ship-evacuation analysis. *Computer-Aided Des.* **2010**, *42*, 1045–1058. [\[CrossRef\]](#)
45. Marín-Morales, J.; Llinares, C.; Guixeres, J.; Alcañiz, M. Emotion recognition in immersive virtual reality: From statistics to affective computing. *Sensors* **2020**, *20*, 5163. [\[CrossRef\]](#) [\[PubMed\]](#)
46. Dziedzickis, A.; Kaklauskas, A.; Bucinskas, V. Human emotion recognition: Review of sensors and methods. *Sensors* **2020**, *20*, 592. [\[CrossRef\]](#)
47. Zhang, H.; Zhang, Y.; Xu, T.; Zhou, Y. Effects of VR instructional approaches and textual cues on performance, cognitive load, and learning experience. *Educ. Technol. Res. Dev.* **2023**, *72*, 585–607. [\[CrossRef\]](#)

48. Vontzalidis, G.; Mystakidis, S.; Christopoulos, A.; Moustakas, K. Spatial Audio Cues in an Immersive Virtual Reality STEM Escape Room Game: A Comparative Study. In *Immersive Learning Research Network. iLRN 2024*; Krüger, J.M., Pedrosa, D., Beck, D., Bourguet, M.L., Dengel, A., Ghannam, R., Miller, A., Peña-Rios, A., Richter, J., Eds.; Communications in Computer and Information Science; Springer: Cham, Switzerland, 2025; Volume 2271. [\[CrossRef\]](#)
49. Jing, A.; Teo, T.; McDade, J.; Zhang, C.; Wang, Y.; Mitrofan, A. Superpowering Emotion Through Multimodal Cues in Collaborative VR. In Proceedings of the 2024 IEEE International Symposium on Mixed and Augmented Reality (ISMAR), Bellevue, WA, USA, 21–25 October 2024; pp. 160–169. [\[CrossRef\]](#)
50. Jinan, U.A.; Heidarikohol, N.; Borst, C.W.; Billinghamurst, M.; Jung, S. A systematic review of using immersive technologies for empathic computing from 2000 to 2024. *Empath. Comput.* **2025**, *1*, 202501. [\[CrossRef\]](#)
51. Saneiro, M.; Santos, O.C.; Salmeron-Majadas, S.; Boticario, J.G. Towards emotion detection in educational scenarios from facial expressions and body movements through multimodal approaches. *Sci. World J.* **2014**, *2014*, 484873. [\[CrossRef\]](#)
52. Ortmann, T.; Wang, Q.; Putzar, L. Facial emotion recognition in immersive virtual reality: A systematic literature review. In Proceedings of the 16th International Conference on Pervasive Technologies Related to Assistive Environments (PETRA '23), New York, NY, USA, 5–7 July 2023; pp. 77–82. [\[CrossRef\]](#)
53. Joo, J.-H.; Han, S.-H.; Park, I.; Chung, T.-S. Immersive emotion analysis in VR environments: A sensor-based approach to prevent distortion. *Electronics* **2024**, *13*, 1494. [\[CrossRef\]](#)
54. XRHealth. Virtual Reality Therapy for Professionals and Patients. *XRHealth*, 2024. Available online: <https://www.xr.health> (accessed on 1 April 2025)
55. Valarmathi, R.; Yazhini, G.; Narmadha, L.S.; Janani, A.M. Enhancing mental well-being through AI-driven virtual environments: The Merrytopia project. In Proceedings of the 2024 International Conference on Power, Energy, Control and Transmission Systems (ICPECTS), Chennai, India, 8–9 October 2024; pp. 1–6. [\[CrossRef\]](#)
56. Guo, H.; Gao, W. Metaverse-powered experiential situational English-teaching design: An emotion-based analysis method. *Front. Psychol.* **2022**, *2022*, 859159. [\[CrossRef\]](#)
57. Dubovi, I. Cognitive and emotional engagement while learning with VR: The perspective of multimodal methodology. *Comput. Educ.* **2022**, *2022*, 104495. [\[CrossRef\]](#)
58. Meta. Meta Quest Pro Official Website. 2024. Available online: <https://www.meta.com/es/quest/quest-pro> (accessed on 1 April 2025).
59. Microsoft. HoloLens Official Web Page. 2025. Available online: <https://www.hololens.com> (accessed on 1 April 2025).
60. Vive. Vive Focus 3 Eye Tracker Official Website. 2025. Available online: <https://business.vive.com/eu/product/vive-focus-3-eye-tracker> (accessed on 1 April 2025).
61. Vive. Vive Facial Tracker Official Website. 2025. Available online: <https://www.vive.com/eu/accessory/facial-tracker> (accessed on 1 April 2025).
62. Teslasuit. Teslaglove Official Website. 2025. Available online: <https://teslasuit.io/products/teslaglove> (accessed on 1 April 2025).
63. Sony. Sony's Mocopi Official Website. 2025. Available online: <https://electronics.sony.com/more/mocopi/all-mocopi/p/qmss1-uscx> (accessed on 1 April 2025).
64. Shiftall. HaritoraX Official Website. 2025. Available online: <https://en.shiftall.net/products/haritorax2> (accessed on 1 April 2025).
65. Virtuix. Virtuix Omni One Official Website. 2025. Available online: <https://omni.virtuix.com> (accessed on 1 April 2025).
66. bHaptics. bHaptics TactSuit X16 Official Website. 2025. Available online: <https://www.bhaptics.com/tactsuit/tactsuit-x16> (accessed on 1 April 2025).
67. Dionisio, J.D.N.; Burns, W.G., III; Gilbert, R. 3D virtual worlds and the metaverse: Current status and future possibilities. *ACM Comput. Surv.* **2013**, *45*, 34. [\[CrossRef\]](#)
68. Frey, D.; Royan, J.; Piegay, R.; Kermarrec, A.-M.; Anceaume, E.; Fessant, F.L. Solipsis: A decentralized architecture for virtual environments. In Proceedings of the 1st International Workshop on Massively Multiuser Virtual Environments, Reno, NV, USA, 8–12 March 2008. Available online: <https://hal.inria.fr/inria-00337057> (accessed on 1 April 2025).
69. Beltrán, F.; White, D. What happens in the avatar stays in the avatar. In Proceedings of the 2023 IEEE International Conference on Metaverse Computing, Networking and Applications (MetaCom), Kyoto, Japan, 26–28 June 2023; pp. 780–785. [\[CrossRef\]](#)
70. Hurst, W.; Spyrou, O.; Tekinerdogan, B.; Krampe, C. Digital art and the metaverse: Benefits and challenges. *Future Internet* **2023**, *15*, 188. [\[CrossRef\]](#)
71. The Metaverse Standards Forum. Available online: <https://metaverse-standards.org/> (accessed on 1 April 2025).
72. IEEE Std 3812.1-2023; IEEE Standard for General Requirements for Identity Framework for Metaverse. IEEE Standards Association: Piscataway, NJ, USA, 2023; pp.1–25. [\[CrossRef\]](#)
73. Fernández-Caramés, T.M.; Fraga-Lamas, P.; Suárez-Albela, M.; Vilar-Montesinos, M. A fog computing and cloudlet based augmented reality system for the industry 4.0 shipyard. *Sensors* **2018**, *18*, 1798. [\[CrossRef\]](#)
74. Ventikos, N.P. Exploring fire incidents/accidents onboard cruise and passenger ships. *SPOUDAI-J. Econ. Bus.* **2013**, *63*, 146–157.

75. Acejo, I.; Sampson, H.; Turgo, N.; Ellis, N.; Tang, L. *The Causes of Maritime Accidents in the Period 2002–2016*; Seafarers International Research Centre (SIRC): Cardiff, UK, 2018.
76. NAFTES. Fire Safety Operations Booklet SOLAS Reg.16.2. 2024. Available online: <https://etraining.naftes.eu/courses/LIB119/> (accessed on 1 April 2025).
77. Koivisto, J.; Hamari, J. The rise of motivational information systems: A review of gamification research. *Int. J. Inf. Manag.* **2019**, *45*, 191–210. [CrossRef]
78. Angra, S.; Sharma, B.; Sharma, A. Analysis of Virtual Reality and Augmented Reality SDK's and Game Engines: A Comparison. In Proceedings of the 2022 International Conference on Edge Computing and Applications (ICECAA), Tamilnadu, India, 13–15 October 2022; pp. 1681–1684.
79. Anderson, M. *Putting the “Reality” in Virtual Reality: New Advances through Game Engine Technology*; Dr. Rudolf Habelt GmbH: Bonn, Germany, 2008.
80. The Khronos Group. OpenXR Overview-The Khronos Group Inc. 2025. Available online: https://www.khronos.org/api/index_2017/openxr (accessed on 1 April 2025).
81. Blender. Blender Documentation. 2025. Available online: <https://docs.blender.org/> (accessed on 1 April 2025).
82. Ship-Broker EU. GENERAL CARGO SHIP (TWEEN DECKER)-2011 Korea. 2025. Available online: <https://www.ship-broker.eu/general-cargo-ship-tween-decker-2011-korea/> (accessed on 1 April 2025).
83. Bergroth, L.; Hakonen, H.; Raita, T. A survey of longest common subsequence algorithms. In Proceedings of the Seventh International Symposium on String Processing and Information Retrieval. SPIRE 2000, A Coruña, Spain, 27–29 September 2000; pp. 39–48. [CrossRef]
84. Coco, M.; Keller, F. Scan Patterns Predict Sentence Production in the Cross-Modal Processing of Visual Scenes. *Cogn. Sci.* **2012**, *36*, 1204–1223. [CrossRef]
85. Koç, C.K. Analysis of sliding window techniques for exponentiation. *Comput. Math. Appl.* **1995**, *30*, 17–24. [CrossRef]
86. Ekman, P.; Friesen, W.V. *Facial Action Coding System: A Technique for the Measurement of Facial Movement*; Consulting Psychologists Press: Palo Alto, CA, USA, 1978.
87. Krumhuber, E.G.; Skora, L.I.; Hill, H.C.; Lander, K. The role of facial movements in emotion recognition. *Nat. Rev. Psychol.* **2023**, *2*, 283–296. [CrossRef]
88. Lessiter, J.; Freeman, J.; Keogh, E.; Davidoff, J. A cross-media presence questionnaire: The ITC-Sense of Presence Inventory. *Presence* **2001**, *10*, 282–297. [CrossRef]
89. Lindquist, K.A.; Jackson, J.C.; Leshin, J.; Satpute, A.B.; Gendron, M. The cultural evolution of emotion. *Nat. Rev. Psychol.* **2022**, *1*, 669–681. [CrossRef]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.