Enhancing sand table based incident command training with extended reality and interactive simulations: A use case in forest firefighting

(Article begins on next page)

Enhancing Sand Table-based Incident Command Training with Extended Reality and Interactive Simulations: A Use Case in Forest Firefighting

Lorenzo Valente, *Student Member, IEEE*, Federico De Lorenzis, *Member, IEEE*, Davide Calandra, *Member, IEEE*, Fabrizio Lamberti, *Senior Member, IEEE*

***Abstract*—In recent years, First Responders (FRs) have faced increasing challenges in their operations, highlighting a growing need for specialized and comprehensive training. In particular, the firefighting Incident Commanders (ICs) are playing a pivotal role, providing directions to field operators and making critical decisions in emergency situations. Over time, traditional training tools in this field have evolved, reaching their pinnacle with Augmented Sand Tables (ASTs). ASTs build on Spatial Augmented Reality (SAR), a form of eXtended Reality (XR) that utilizes projections. Although ASTs enable large-scale visualization of the morphological features of the terrain, by relying solely on SAR it is not possible to fully leverage the potential of XR, which is increasingly recognized as a powerful tool for training.**

This work introduces a novel approach to training ICs by integrating ASTs with XR, incorporating a learning-by-doing methodology alongside an objective measurement of trainees' performance. To this end, an XR Training System (XRTS) has been developed, combining the capabilities of an AST with personal Mixed Reality (MR) devices and integrating a physically accurate, interactive fire simulator. This system was deployed within a forest firefighting IC training course. All the system components were designed based on the theoretical foundations of decision-making to effectively develop the necessary skills.

The proposed approach was compared with traditional AST-based training methods for these roles, focusing on the analysis of learning outcomes, user experience, usability, and cognitive load. The study demonstrated several advantages associated with the use of the XRTS, including improvements in training effectiveness and a notable reduction in overall cognitive load.

***Index Terms*—Extended Reality, Training, Sand Table, Simulation, First Responder, Incident Commander, Forest Firefighting, Decision-Making, Cognitive Load.**

I. Introduction

First Responders (FRs), e.g., firefighters, emergency medical services, and police officers, are increasingly important and beneficial figures which have faced increasing operational challenges over time [1].

For severe events, Incident Commanders (ICs) become crucial [2]. These figures are responsible for leading the various FR groups, taking more onerous and vital decisions [3]. In particular, ICs must quickly assess situations, develop risk-based strategies for large-scale incidents, and promptly implement bespoke plans to resolve them [4]. Additionally, the Incident Command System (ICS) was defined as a standardized mechanism to enhance cooperation and communication between various departments and organizations [5].

This work has been supported by the VR@POLITO initiative.

As high-risk situations become more and more frequent, ICs specifically need to develop the appropriate skills to manage FR teams in stressful environments, where human error can lead to fatalities and injuries, due to stress itself and lack of preparedness [2]. However, traditional courses are no longer sufficient to provide the training necessary to achieve the required level of proficiency in an efficient and safe way [6].

Traditional IC training courses have historically emphasized scenario visualization, which is critical for training, preparedness planning, and response management in public safety [7]. Initially, static topographic maps were used, but their limitations drove the development of three-dimensional plastic maps incorporating projected topographic data [8]. This approach aimed to simulate traditional sand tables used for strategic planning and for visualizing terrain features [9]. However, the poor flexibility of these solutions has kept sand tables popular, often augmented with overlays of topographic and satellite maps [10]. This augmentation method, known as Spatial Augmented Reality (SAR) [11], is a particular kind of Augmented Reality (AR) technology that employs projection mapping to superimpose virtual images onto real objects, eliminating the need for traditional screens as well as of Head-Mounted Displays (HMDs) or handheld devices typically used in other AR setups [12]. Such evolved versions of sand tables are commonly referred to as Augmented Sand Tables (ASTs).

SAR, being a form of AR, falls within the broader spectrum of eXtended Reality (XR). A defining feature of XR is interactivity. However, current ASTs focus solely on visualization [13]. The scarce use of interaction suggests that the full potential of XR technology has not yet been fully harnessed in this context. XR is an umbrella term that includes not only AR but also Virtual Reality (VR) and Mixed Reality (MR), which has increasingly gained traction in various fields, especially in immersive applications for education and training [14], [15], [16]. Evidence shows that XR, particularly in the form of XR Training Systems (XRTSs), is highly effective for these purposes, especially in emergency management [1], [17].

Haskins et al. [6] assert that effective XRTSs must offer interactive experiences with physically accurate feedback. Utilizing physics-based simulators, in particular, can enhance realism and training efficacy [18], aiding in the development of decision-making and problem-solving skills [19]. In this context, given the risks of live-fire training, fire simulation is a well-researched application [18], [20], enabling the creation of controlled firefighting scenarios [21].

Moreover, the growing complexity of wildfires demands that ICs develop effective strategies to ensure containment and safety. Currently, firefighting culture is shifting from a proactive approach to a more defensive strategy, emphasizing risk analysis of assets [22]. Specifically, wildfire management strategies focus on dealing with Decision Points (DPs) [23], which are critical moments when alternative decisions or courses of action (e.g., fire suppression or broader fire management operations) could have been chosen or taken during a fire event. These activities are referred to as maneuvers in the literature [22], and can encompass either a single intervention or a coordinated set of firefighting operations. Under extreme time pressure, evaluating multiple options simultaneously becomes nearly impossible. To address this issue, several serial option evaluation models have been developed, including the Recognition-Primed Decision (RPD) model [23]. This model is based on recognizing a situation and analyzing it by carrying out a comparison with a mental representation of a familiar scenario. The analysis focuses on key factors such as DPs, perceptual cues that facilitate recognizing and interpreting environmental signals (e.g., fire intensity), and causal factors. The possible response strategy is then evaluated for plausibility and feasibility, often through mental simulation of its potential consequences. Depending on the assessment, the strategy may be implemented directly, adapted to better suit the situation, or discarded if found unsuitable, with a new strategy aligned with the identified prototype subsequently evaluated.

Currently, the standard training for ICs within forest firefighting relies on ASTs, employed to support formal courses for the command roles. An example comes from the Italian forest firefighting corps of the Piedmont Region, Italy, namely *Corpo Volontari Antincendi Boschivi (AIB) del Piemonte*[1] (in the following abbreviated *AIB Corps*). Concepts tackled in this course are applied practically applied using an AST configured to represent specific territories and wildfire scenarios. Trainees, equipped with real management tools (e.g., maps, weather data, markers), make tactical and strategic plans for wildfire management. An instructor validates these plans based on his or her experience and actual fire behavior. The AST provides a shared view of the scenario, allowing trainees to collaboratively apply theoretical knowledge to manage the case study.

Even though it represents a cutting-edge training approach for ICs in firefighting, the course above is affected by some limitations due to the technology used (AST) and the high cognitive load it puts on trainees. A significant drawback is the need for an instructor to subjectively evaluate the trainees' decisions, since AST visualizations cannot replicate the effects of real-world maneuvers. Although wildfire propagation simulators could theoretically enhance training, they typically do not account for firefighting maneuvers that could alter simulation outcomes.

The proposed work aims to introduce a new paradigm for the training of forest firefighting IC figures that can overcome the mentioned limitations by leveraging the benefits of XRTSs, which have been already demonstrated for field operators (e.g., squad members) directly engaged in related activities [24]. Specifically, the study was driven by the following hypotheses:

- **H1:** the use of an XRTS fosters decision-making skills on large-scale firefighting scenarios, thus improving the assimilation of the procedural details of the considered tasks as well as enabling the trainees to correctly manage the AIB Corps maneuvers and schedule interventions better than with the traditional AST-based training.
- **H2:** the use of an XRTS has a lower impact in terms of overall cognitive load perceived by the targeted figures during the training experience.

The hypotheses were tested using both objective and subjective measures on a new XRTS, which is based on an AST but enhanced through two main modifications. First, it integrates a state-of-the-art, open-source cell-based fire propagation simulator (*Cell2Fire* [25]), modified to support all firefighting interventions relevant to the considered command roles and to provide an objective visualization of the results produced by the devised maneuver. Second, an MR handheld application serves as a personal interface to the XRTS, enabling the visualization of 3D information over the sand table and allowing each trainee to have an individual instance of the initial case study projected onto the AST. This setup enables independent work, using the handheld device as a "window" into the simulation, superimposed as MR content onto the physical table.

Trainees can use the simulator to practice and refine their decisions until they reach satisfactory results through learning-by-doing. The simulation results can be projected onto the AST to facilitate the eventual debriefing by the instructor.

In order to evaluate the benefits of this approach, a user study was conducted comparing the use of the proposed XRTS with a traditional training approach based solely on AST. The XRTS was developed in collaboration with the AIB Corps to ensure its relevance and consistency, and its effectiveness was evaluated within a training course designed for novice team leaders of the mentioned firefighting body [21].

The contributions of this study can be summarized as follows. First, the study demonstrated that the XRTS improved understanding of the firefighting scenario, enabling better strategic decision-making, which confirmed the first hypothesis (H1). Second, it showed that the XRTS reduced perceived workload, stress, and task complexity, while enhancing time efficiency, thus supporting the second hypothesis (H2).

II. Background

In the following, a comprehensive review of relevant literature will be presented, alongside an outline of the research gaps considered in the design of the devised XRTS.

A. XR-based Training for First Responders (FRs)

Recent literature extensively explored the use of XR technology to develop training tools for FRs, focusing on both VR [6], [18] and AR/MR [16], [17], [26].

Moreover, the possibility to integrate Virtual Environments (VEs) with more or less realistic representations of hazardous

[1] AIB Corps: https://corpoaibpiemonte.it

elements, such as fire and flames, has led to numerous studies in the field of firefighting.

Chen et al. [27] presented an XRTS designed to enhance situational awareness during building fire emergencies. By combining Building Information Modeling, Internet of Things (IoT), and VR/AR technology, the system was able to create an immersive VE, also integrated with realistic fire behavior. To this purpose, instead of relying on mathematical simulations, the visualization of smoke and fire was driven by pre-recorded measurements from IoT sensor data (i.e., heat and smoke detectors). This approach reduced trainees' psychological pressure and improved evacuation performance, since the realistic behavior of fire enhanced their situational awareness.

A similar approach was followed by Calandra et al. [24], who proposed a VR Training System (VRTS) aimed at training beginner forest firefighters in the use of three firefighting hand tools. The system integrated a non-physically-based yet plausible cell-based spreading logic for fire simulation. The effectiveness of the VRTS was assessed by incorporating it as supplementary training within a formal course on the subject and by comparing final outcomes, including theoretical and practical exam results, with those of participants who underwent only traditional training. Despite the simplified simulation, the system proved to be capable of enhancing procedural learning, motivation, and perceived training quality compared to traditional methods.

What these and other prior studies regarding the use of XRTSs in the firefighting context have in common is that they are mainly focused on scenarios involving field operators directly engaging with fires, leveraging the high interaction fidelity of immersive VR [24]. However, higher responsibility figures commanding squads and overseeing tactical and strategic planning are typically not investigated. Nonetheless, XR technology has significant potential to enhance the training of these command figures, whose needs differ from those of fireground operators.

B. XR for Incident Commanders (ICs)

ICs require a privileged viewpoint of the scenario, enabling them to engage in both strategic planning and decision-making activities. In these contexts, training tools conventionally used include maps, models, and traditional sand tables. At the same time, XR-based tools are extensively employed by FR ICs to visualize geographic data spatially, aiding in training, analysis, or planning scenarios.

For instance, Woods et al. [10] presented a pioneering study on an AST to enhance the spatial thinking abilities of physical geology students at East Carolina University. The proposed AST integrates a sand table with virtual contour lines and a water flow model, created using a 3D scanning camera, visualization software, and a projector. This virtual water flow model illustrates the dynamics of water across the surface. The authors showed that, compared to the traditional course, the AST significantly improved spatial perception of the environment, enhancing understanding of teaching objectives.

The benefits of ASTs for training ICs have also been studied in the context of FR tasks. As a matter of example, Amburn et al. [28] introduced *ARES*, an AST that utilizes readily accessible commercial products to create a cost-effective method for visualizing geospatial terrain, also incorporating a tangible UI specifically tailored for simulation and training applications targeted to the U.S. Army.

Abich et al. [29] conducted a preliminary study on the impact of *ARES* on training in the context of a CBRN course for ICs. They divided trainees into two groups: one used *ARES* and the other used traditional multimedia content. The evaluation included course assessments, knowledge tests, and self-reported questionnaires on teamwork, self-efficacy, and technology acceptance. Results showed a favorable perception of *ARES* in enhancing course outcomes and improving cohesion in collaborative tasks, though no significant differences in learning gain were actually observed.

When the tangible component of ASTs is not essential but the benefits of visualization are desired, fully virtual solutions can be used. Korkiakoski et al. [30] introduced *ARTS* (AR Tactical Sandbox), an MR tool for providing real-time immersive analyses of mission-critical information through a virtual AST, visualized as 3D content via Microsoft HoloLens 2. The tool aimed to enhance situational awareness in tasks like locating and rescuing individuals from dangerous areas. Interestingly, they found that single-user training in MR outperformed multi-user training in terms of situational awareness, contrary to the expectation that collaboration could enhance overall performance.

Mao and Chen [31] presented an XRTS centered on the concept of Common Operational Picture (COP), a shared centralized visualization that provides a comprehensive and up-to-date overview of a given situation to all stakeholders involved in operations, integrating real-time intelligence and accurate positioning. Their XRTS utilized handheld MR devices, such as tablets and smartphones, to enhance the COP by reconstructing the environment in 3D, using overlays, symbols, and animations. The aim was to enhance tactical training effectiveness for Taiwanese Army FRs, improving understanding of emergency procedures and enhancing tactical plan management directly within operational environments. A comparative analysis of the devised tool against traditional COP-based courses showed its superior outcomes in these aspects.

As it can be inferred from the last two works, unlike solutions based only on SAR (like ASTs), the use of MR also allows for adding content that can be positioned above and around the sand table itself, such as 3D indicators, information panels, or virtual elements related to the displayed use case.

C. ASTs for Firefighter Training

An extremely relevant application of ASTs in the field of firefighting is *SimTable*[2], firstly presented by Guerin and Carrera [32]. The original system utilized an AST to visualize the spread of forest fires, primarily serving as a training tool for ICs in forest firefighting with a focus on command and tactical roles. In the latest iteration, the simulation engine, currently branded as *AnyAzard*, has been enhanced to support a

[2]SimTable: https://www.simtable.com/

variety of alternative scenarios, including flooding, simulation of vehicular traffic, and the dispersion of smoke or chemical plumes.

Focusing on the fire scenario, the platform allows trainers to choose from eight known wildfires and their corresponding topographies. After selecting a wildfire, the platform displays a color-coded terrain map that guides users on where to add or remove sand to accurately replicate the territory. GIS layers, including vegetation, roads, and buildings, can be interactively added to influence fire behavior. The fire behavior is not managed through a physically accurate simulation, since the system employs a custom Cellular Automata fire spreading model in which the fire spreads probabilistically to neighboring cells, with rates determined by external matrices drawn from the BehavePlus [33] fire modeling system. Factors like wind, elevation, and fuel type affect ignition probability, with cells downwind and uphill from a burning cell being more likely to ignite. Fires can be initiated using a real lighter or a laser pointer on the table. From a training perspective, the system provides valuable information, such as wind direction and time data, to facilitate collaborative analysis of selected case studies. It allows for two types of firefighting interventions: creating firebreaks (clearing fire lines along wildfire paths) and using aircraft for indirect attacks (dropping fire retardants at targeted locations). However, maneuvers only render cells non-burnable or delay ignition without detailed physical modeling. The system lacks support for interventions directly on the fire front and does not assess intervention feasibility, efficiency, or safety. Despite enabling multiple trainees to be co-located, it manages only one simulation instance projected onto the table, which prevents its use in scenarios requiring the simultaneous participation of multiple operators.

D. Contributions

Based on this review of the state of the art, it was possible to identify the minimum set of interventions that a training system based on an AST should offer in order to be applicable to IC training in firefighting contexts. In parallel, an ATS actually used in a course targeted to firefighting operators training was studied to analyze its strengths and weaknesses, aiming to maximize the potential contribution of XR technology.

As a result, an XRTS centered on an AST was proposed and collaboratively developed with the AIB Corps to support the training of IC figures involved in forest firefighting.

To address the limitations of previous fire spreading models [32], it was decided to use *Cell2Fire*, a physically accurate cell-based fire simulator [25]. This tool is based on Canadian Forest Fire Behaviour Prediction [34] and on Prometheus [35], a widely used wildfire spreading model. Certain modifications were necessary to let it communicate with other components of the XRTS and introduce the necessary functionalities for the studied context. Specifically, the simulation was upgraded to start from any fire state, incorporating data from actual wildfires (originally, it could only start from single ignition points). Furthermore, all the aspects of forest firefighting interventions management have been meticulously modeled based on guidelines provided by the AIB Corps. The modified version of the simulator is detailed in Section IV.

To mitigate the issues related to the two-dimensional nature of projected content, it was decided to introduce additional MR handheld devices (i.e., tablets) alongside the AST, as done in [31]. These devices are provided to trainees and serve both as interfaces to the system (via touchscreen) and, as said, as sort of windows onto the virtual world. In particular, they enable overlaying a virtual 3D representation of simulated content onto the AST, leveraging MR functionalities such as marker detection and SLAM (Simultaneous Localization and Mapping). The decision to use handheld MR technology instead of HMDs was also made to mitigate potential barrier issues related to the technology, which are often present in such solutions [21].

The limitation of having a single instance of the simulation displayed on the table and shared among all the trainees was addressed using the tablets. Through the tablets, the trainees can align themselves with the case study set up by the instructor on the AST. Then, from this point on, they proceed with a local version of the simulation visible only through their device in MR (superimposed to the AST content), allowing them to work independently while others do the same. This feature supports learning-by-doing through self-directed training sessions where trainees can test their own firefighting strategies, simulate the effects, and adjust their choices until achieving a satisfactory outcome. The final result can then be shared with the instructor for collective viewing on the AST, e.g., during debriefing sessions.

Finally, to cope with the mentioned lack of proper validation of firefighting training based on immersive technologies [21], the XRTS was evaluated within an official course targeted to ICs which is organized by the AIB Corps. This approach draws inspiration from a previous study [24] that focused on the use of an XRTS to train novice FRs without command roles.

III. Current Use of the AST by the FR Body

As mentioned in Section I, the proposed XRTS builds upon an AST currently utilized for forest firefighting training. This AST is situated at the FORMONT[3] training center of the AIB Corps in Peveragno, Italy and serves as a focal point for training IC roles within the AIB Corps. The IC figures trained through this tool are essentially three. Team leaders use the AST for tactical planning, wildfire prediction, and resource management. AIB Coordinators (CoAIB) focus on identifying strategic points and allocating resources. Analysts develop comprehensive wildfire strategies using the AST to forecast fire progression and identify effective interventions.

For each of these figures, the AST is a key component of (different) formal courses focused on teaching procedural and theoretical concepts on territory morphology and standard firefighting interventions. These courses cover topics such as cartographic frameworks for strategic planning, forest fire regulations, AIB Corps interventions, and risk analysis for deployment and operational effectiveness.

The current study focuses on the IC role at the entry level of the hierarchy, i.e., on team leaders. While subjects training in this role possess basic knowledge and some fire containment

[3]FORMONT: https://www.formont.it/formont/centri/peveragno/

experience, they lack a broader strategic perspective and field experience. To earn the certification, they must learn to accurately predict wildfire spread and assign the appropriate operational squads. Their goal is to defend key strategic points while effectively managing available resources. The training also covers optimal deployment ranges for crews and vehicles to ensure safety and operational efficiency.

A. *Traditional Training with the AST*

The standard training consists of multi-day, theoretical courses with lectures (mainly leveraging video content), followed by a practical session on the AST to reinforce understanding of the learned concepts. After completing the course, the trainees take an exam to obtain the certification. Each course session can accommodate approximately 30 trainees.

The practical session on the AST is preceded by a theoretical lesson conducted using the AST itself (already configured for a specific use case), in which the territory morphology, including critical elements like the presence of key sites or targets such as homes, cities, or power lines, is explained. Additionally, the instructor shows the progression of a past wildfire step-by-step using 2D areas highlighted in red. These visualizations are generated using aerial photographs captured during the wildfire, supplemented by satellite data.

After the theoretical lesson, the trainees are introduced to the use of the AST for the practical session. The instructor explains the techniques and tools available with the AST, and provides detailed cues on the roleplay, ensuring that trainees are well-prepared for their specific responsibilities. For team leaders, for instance, the instructor acts as a CoAIB serving as one of the directors of extinguishing operations (*Direzione delle Operazioni di Spegnimento*, DOS), assigning to each trainee a strategic objective (i.e., a specific location or area) to manage and a set of available resources, including personnel and vehicles.

The practical session then consists of self-directed training, usually conducted in pairs, in which the trainees are positioned around the table. Each pair can discuss the best way to protect the assigned objective (usually displayed on the AST through physical flags, Fig. 1a), applying concepts learned during the course and aiming to design the best tactical or strategic plan for managing the specific area of the case study they are working onto. To compute the expansion of the fire front, the trainee, based on Campbell's expansion model [36], can apply the so-called "vector method", by placing on the table various types of vectors with three different intensities (i.e., low, medium, strong). The three vectors are individually parameterized: the wind force, the slope of each cell, and the intensity of each cell, which results from the combined effect of the fuel and solar exposure at a given time instant. After placing the three vectors, the trainee can calculate the vector summation of the future expansion and then proceed with the definition of the maneuver. It should be noted that the effort required to calculate the expansion of the fire front at each time instant is not a negligible factor, requiring meticulous attention to avoid inaccuracies. This task is both cumbersome and repetitive, possibly placing a heavy cognitive load on those performing it.

Another action performed on the AST is to indicate, using again markers like flags, the positions where interventions are planned. To validate the feasibility of these planned maneuvers and estimate their effects on the fire progression, the trainee can consult the instructor. Upon hearing a detailed explanation of the drafted maneuver (tactical or strategic), the instructor provides feedback based on his or her knowledge and experience. Therefore, the outcome heavily relies on subjective factors.

The practice session is typically followed by a debriefing phase, in which the instructor uses the AST, with all the trainees around it, to collectively analyze all the tactical or strategic plans drafted during the self-directed practice. This includes correcting any errors, providing alternative solutions, or demonstrating how the fire progressed in the real-life scenario. Also in this case, while the instructor's evaluation is valuable due to his or her field experience and understanding of the case study, the subjective nature of the assessment remains a significant factor to consider.

Based on the analysis of the characteristics of this AST-based training, combined with the review of existing solutions in the state of the art, it was possible to identify the set of functionalities required for the XRTS being designed to overcome current limitations of IC firefighting training and fully leverage the potential of XR technology.

IV. XR Training System (XRTS)

As explained earlier, the devised XRTS has been developed to work in conjunction with and around the AST, in order to make it possible to seamlessly integrate it into existing training protocols based on the AST usage. Its elements were designed based on the RPD model [23], to enhance trainees' decision-making skills while also addressing the specific need of the AIB Corps to improve training effectiveness and achieve outcomes comparable to traditional methods. Some relevant elements are shown in Fig. 1. As it can be seen from the system architecture reported in Fig. 2, the proposed XRTS consists of four main components: the sand table augmented using SAR, the modified version of *Cell2Fire*, the SAR and Fire Simulation server, and the MR tablet application.

A. *Modified Cell2Fire*

Concerning fire simulation, it was chosen to opt for a physically accurate, cell-based logic. *Cell2Fire* [25], an open-source raster-grid fire simulator, was selected as a starting point and customized to meet the requirements of standard forest firefighting training. As said, the simulator drives the fire life cycle and its rate of spread through the use of the Canadian Forest Fire Behavior Prediction system [34] combined with Prometheus [35], a simulation model built on Hyugens' principle of wave propagation [37].

Moreover, the simulator relies on detailed morphological and meteorological data to model the fire dynamics. At the cell level, it considers factors like fuel type, slope, and elevation. Globally, it incorporates meteorological inputs such as wind speed, wind direction, and relative humidity. During simulation, each cell on the terrain is assigned a status that

(a) Objectives on the AST (SAR) (b) Objective and interventions (MR) (c) Fire intensity & fuel tooltip (MR) (d) Wind sock (MR) (e) Smoke visualization (MR)

Fig. 1. Relevant elements of the proposed AST-based XRTS.

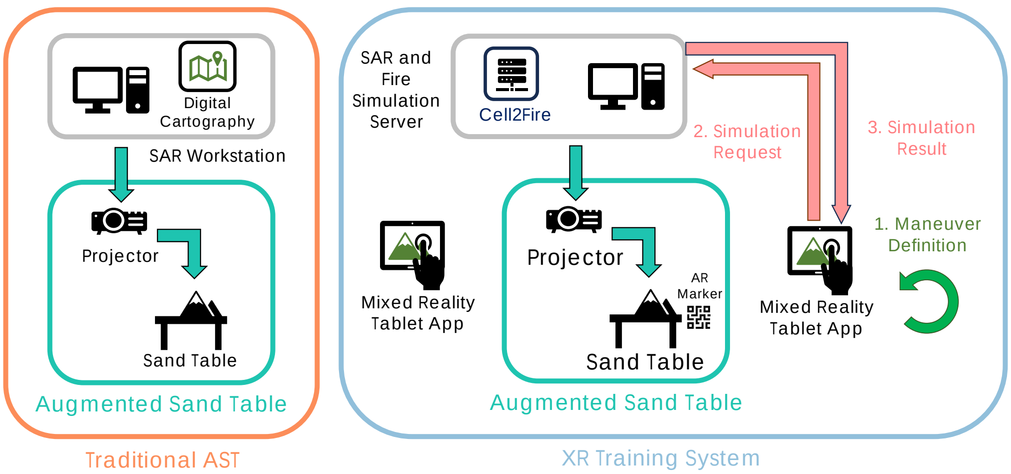


Fig. 2. Traditional AST used in the considered context (left) and architecture of the devised AST-based XRTS (right).

evolves over time. Initially, cells are categorized as Available (with fuel), Harvested (fuel removed), or No Fuel (non-burnable due to features like roads or water). As the simulation progresses, additional statuses like Burning (actively burning with sufficient fuel) and Burnt (fuel exhausted or no longer influencing fire spread) are applied. Burnt cells transition to No Fuel status and cease to contribute to fire spread calculations.

With *Cell2Fire* it is possible to obtain the new fire expansion fronts starting from an initial situation of the case study and a set of custom ignition points. Hence, the first modification required was the introduction of the capability for the simulator to start from a specific fire state provided as input, enabling required calculations to be performed starting from it rather than just from individual ignition points on single cells. This functionality was essential in order to let the instructor run simulations based on past fire data.

Furthermore, to enable the possibility for *Cell2Fire* to manage firefighters' interventions, a dedicated logic was integrated in the simulator too. This logic models the various interventions needed by forest firefighting ICs for effective wildfire management, which can be categorized into three main types: interventions that indirectly limit the spread of the fire, such as firebreaks, which convert the state of a cell from Available to Harvested, thereby blocking the future expansion of the fire in that area; interventions that release water on Burning or Available cells, such as air tankers, which lower the fire intensity value in the selected cell only during release, reducing the probability of fire expansion from that cell; interventions that use fire, such as tactical fire, which converts the status of a cell from Available to Burning, inducing a new fire.

The interventions, detailed in Table I, vary in their characteristics, and the choice of the one(s) to use depends on factors such as fire intensity, available resources, and territory morphology. Each type of intervention requires specific resources, crucial for implementing an effective maneuver. For instance, aircraft attacks require considerations about the number of available vehicles and supply points, whereas direct attacks involve factors such as safe distances for engagement and supply methods. Given that wildfire management operates on hourly time steps (with maneuvers assumed to start and finish within a single step), each planned intervention is designed to affect the simulation for one hour from the time of insertion.

The simulations are executed through a Dockerized image of *Cell2Fire*, which is run by a server application.

B. Spatial AR (SAR) and Fire Simulation Server

This server application, developed in Unity 2020.3 and running on a workstation connected to the projector, manages content projection on the AST, executes fire simulations, and interacts with MR clients running on tablets. Through a UI

TABLE I
FIREFIGHTING INTERVENTIONS SUPPORTED BY THE DEVISED XRTS. THE SYMBOL ✓ INDICATES INTERVENTIONS THAT ARE FULLY AVAILABLE (RESPECTING AVAILABILITY OF RESOURCES), ✓* PARTIALLY AVAILABLE (CAN BE REQUESTED, BUT THEY ARE NOT UNDER THE RESPONSIBILITY OF THE TRAINED IC FIGURES), X NOT AVAILABLE, AND ✔ MOST RELEVANT INTERVENTIONS FOR THE CORRESPONDING IC FIGURE.

Intervention	**Availability**			**Description**
	Team Leaders	**CoAIBs**	**Analysts**	
Direct Attacks				
Handcrew	✔	✓	✓	Operators who use individual tools to extinguish low-intensity fires and clean up the ground. Without their intervention, any type of flame can start a new fire.
Light Truck	✔	✓	✓	Light trucks to supply operators who spray water with high-pressure nozzles directly at the firefront. Refueling can be done by drawing water from a source up to 100 meters away, replacing the empty module, or using a tanker truck with up to 6,000 litres of water.
HPL	✔	✓	✓	High-pressure line to supply operators who spray water with high-pressure nozzles directly at the fire front. Refueling can be done by drawing water from a source up to 100 meters away or using a tanker truck holding 3,000 to 6,000 liters of water.
Aircraft Attacks				
Air tanker	✓*	✔	✓	Airplanes discharge large amounts of water. They replenish their supply only horizontally and need special access paths to water reservoirs. Each release covers two adjacent areas of 50×90 square meters and cannot be modularized.
Helitack Light	✓*	✔	✓	Small regional helicopters equipped with 700-1,000 liter tanks. They can vertically descend to water supply points or artificial tanks to replenish. Each release covers an area of approximately 30×80 square meters.
Helitack Heavy	✓*	✔	✓	Large helicopters equipped with tanks holding up to 9,500 liters. They replenish their water supply like the smaller helicopters but do not use artificial tanks. Each release covers an area of approximately 50×90 square meters.
Ground Attacks				
Firebreak	✔	✓	✓	Removing all fuel in an area to make the land completely unburnable, either manually using blowers, brush cutters, and chainsaws, or with machinery, such as bulldozers.
Fire Attacks				
Tactical Fire	X	X	✔	Controlled fire front at a safe distance from the main fire to isolate an area by consuming fuel. This intervention is used when building a firebreak is not possible or near mountain population centers.
Backfire	X	X	✔	Intentionally igniting a fire at a distance during special emergencies with uncontrollable fires. The new fire is drawn towards the main fire by the depression it generates, causing the two fires to collapse into each other.

navigable via mouse and keyboard, the instructor can configure the case study projected onto the sand table, create "rooms" (acting as starting points for case studies where users connect via tablets), and manage the projected visualization (e.g., navigate the simulation timeline, change map type, select different simulations, etc.).

The visualization includes a 3D digital model related to the case study, reconstructed at runtime using Unity Terrain with data from case study files. Territory morphology comes from a height-map file and textures are derived from topographic maps projected onto the AST. The *TerrainGridSystem* asset for Unity [38] overlays a visual grid on the terrain, indicating details such as cell number and size. Fire intensity data are used to color active burning cells from red to yellow and mark burnt cells in black. Transitions between fire states use linear interpolation of cell activation and fire intensity values to visualize wildfire expansion as well as to determine intervention timings, reflecting smooth transitions of the fire front. The combination of morphological information derived from the AST and fire intensity data was designed to leverage perceptual cues, a critical skill for decision-making. This information supports planning by enabling rapid assessment of the situation at a glance, thereby accelerating the selection of appropriate interventions.

Initially, the case study loads without generating a simulation, allowing interpolation between known fire states. When a simulation is requested through a tablet, the server computes and adds it to the room. It is then possible to visualize it. Communication between the application and the tablets is managed by a Java Spring server, serving as a proxy for calls between the tablets and the Unity server. The server also handles room creation and simulation data saving at the file system level.

C. *Mixed Reality (MR) Tablet Application*

The MR application, developed again using the Unity 2020.3, has been designed to be used via an Android tablet (Samsung Galaxy Tab S7 devices were used). The application takes advantage of the same assets used to manage content projection onto the AST (i.e., the 3D terrain and the overlaying grid), but with two fundamental differences.

Firstly, the 3D terrain is not viewed from above but is anchored to the physical table as MR content, through marker detection (via ARFoundation) for the first placement, and then through SLAM (via Google ARCore) for maintaining it anchored to the AST. Secondly, the grid cells are interactive, allowing the users to select them through the touch interface of the tablet. For instance, the users can display information such as fuel type or fire intensity (via a tooltip on the selected cell, Fig.1c). The application includes a 2D UI that allows the users to configure the visualization (e.g., turn off smoke, cycle through available objectives, navigate forward and backward in the simulation timeline), as well as select and insert the various interventions (Fig. 1b). It also provides other information such as time and weather data (e.g., wind direction and speed, Fig. 1d). All these functionalities were designed to enable

ICs to clearly identify causal factors, thereby accelerating and optimizing the definition of maneuvers.

After inserting the desired interventions on the selected cells, the system automatically sends the data to the simulator, which within seconds returns the updated simulation output for display on the table. Similar to what happens with cells on fire, cells where interventions have been inserted are visually modified to display the corresponding intervention icon.

When applying interventions on the selected cells, the XRTS automatically evaluates the following aspects in descending priority: *safety*, assessing whether the fire intensity poses a risk to operators involved in the intervention; *effectiveness*, determining if the chosen intervention can control wildfire spread; *efficiency*, evaluating whether resource allocation aligns with optimal intervention requirements. These three aspects are used to define the goal to be achieved when managing a DP under conditions of extreme time pressure. Given the lack of time to analyze all the possible options, ICs must prioritize to plan the maneuver as conservatively and defensively as possible. If safety risks are too high, the system may block the insertion of the intervention and alert the user. This feature, which can be seen as a sort of automatic variable calculation for decision-making, was modeled around the type of feedback provided by AIB Corps instructors during the training sessions with the AST. If the XRTS prevents the user from selecting an excessively incorrect intervention, it is mirroring the actions an instructor would take when overseeing a traditional training session with the AST. Errors in effectiveness or efficiency, on the other hand, only reduce the intervention score, which is shown to the user after each insertion.

To encourage a trial-and-error approach, the system allows users to backtrack, modify, or remove interventions, generating new simulations at need. This functionality was designed to enhance decision-making skills, as it is essential for ICs to adapt or redesign their tactics [23].

To display information about the objective, it was decided to use the same visualization paradigm adopted in the traditional AST-based course, with each objective represented by a flag on the digital terrain model in the corresponding cell (Fig. 1b) to ensure that it remains clear and visible for defining tactics.

Regarding smoke visualization (Fig. 1e), the *Ultra Real Fire Effects Volume 1* asset for Unity [39] was used to create a particle system for each burning cell. Each smoke particle rises to varying altitudes controlled by the fire intensity to simulate realistic behavior, as smoke height is crucial for firefighters to understand fire intensity before reaching the site. This visualization can be toggled on and off at runtime due to potential interference with cell visualization and its computational demand when moving through the simulation timeline.

V. Experimental Evaluation

This section describes the user study conducted to evaluate the overall effectiveness of the proposed XRTS. Specifically, a new training protocol based on the developed system was designed and compared to the traditional AST-based forest firefighting course. The study considered as use case the course for novice team leaders organized by the AIB Corps. Team leaders were deemed as representative of the IC category, being allowed to request the majority of interventions supported by the XRTS. At the same time, their responsibility is confined to a small portion of the fire (specifically, the assigned objective), limiting the range of possible strategies. This factor facilitates the assignment of DP-based tasks of similar complexity to all the trainees, enhancing comparability in the evaluation of their performance.

A. Participants

The study involved 32 participants (30 males and 2 females) aged between 20 and 56 ($\overline{x} = 36.62$, $s = 10.60$) who volunteered among the trainees enrolled for the first time in the mentioned IC course. The participants reported limited to no experience with XR technology ($\overline{x} = 1.87$, $s = 0.93$ on a 1-to-5 scale, with one corresponding to “I have never used XR technology” and five to “I use XR technology every day”), and no experience with the sand table.

B. Study Design

The 32 participants were randomly assigned to two groups of 16 subjects each to avoid potential self-selection bias: the *Augmented Sand Table* (AST) group (Fig. 3a), which followed the standard course schedule with a traditional practice session on the AST using the vector method and instructor feedback; the *eXtended Sand Table* (XST) group (Fig. 3b), which experienced an alternative course version where the devised XRTS replaced the AST-only practice session.

The study was driven by hypotheses H1 and H2, presented in Section I.

As explained in Section III, the current training provided to IC figures of the AIB Corps consists of a course with a clear structure, consisting of theoretical and practical components. The course begins with two days of video-based classroom lessons covering territory morphology and strategies for managing large-scale wildfires from a role-specific perspective. These lessons are followed by another half-day lesson in which the instructor introduces the AST and describes the morphological features that can be visualized directly on the territory. In this lesson, trainees only act as spectators, since the focus is on making them familiarize with the AST rather than engaging them in actual training. For this reason, the use of the XRTS was not deemed useful and, consequently, it was not included in the comparison. Then, practical training on the AST spans half a day split into two sessions. In the first session, an instructor provides an introduction, detailing the case study configured on the AST. In the second session, the trainees, divided into pairs, play the role of a team leader arriving at the fire front and being tasked with planning maneuvers for two DPs, spaced one hour apart (representing two simulation steps), to slow down the fire progress and protect a specific assigned objective. The course concludes on the fourth day with an instructor-led debriefing session to review the trainees’ performance, followed by theoretical and practical exams. The former covers concepts beyond the AST; hence, it was not included in the investigation. The

(a) Traditional training with the AST (b) Training with the devised XRTS (c) Practical exam

(d) Vector method (traditional training) (e) MR tablet application (XRTS) (f) Action plan validation via XRTS

Fig. 3. Traditional IC training of AIB Corps with the AST (a, d), training with the proposed AST-based XRTS (b, e), and IC course's practical exam (c, f).

latter, instead, served as the main measure of learning for the study (Fig. 3c). It consists of a simulation of a real wildfire scenario conducted within a (simulated) control room, in the same conditions as the practice session (arrival at the fire front, and maneuvers for the DPs within the next two hours only). The trainees are provided with all the materials that would be available in a real-world scenario, including maps, measurement tools, and potential communication lines with other units simulated by the instructor.

Some videos exemplifying the training with the two approaches, participants' interaction with the MR application, and the practical exam are available online[4].

C. XRTS Preparation

For the experiment, it was necessary to prepare the case study used in the traditional course for incorporation in the devised XRTS. Specifically, during the study period, the AST had been configured with the Limone Piemonte case study, related to the forest fire that occurred in the mountains surrounding that city in the North of Italy on October 30, 2021. To this end, in collaboration with the AIB Corps, all necessary data required for managing the projection and launching the simulations were gathered and provided in input to the XRTS.

The cartographic data, such as maps, fuel layers, roads, rivers, and lakes were manually extracted from GIS and converted into the proper input format. Information regarding weather and the progression of the real, past fire was collected with a granularity of two hours per fire state. Additionally, in accordance with AIB Corps maneuvers, a cell size of 50m $\times$ 50 m was selected for the simulation/terrain grid.

With the assistance of the AIB Corps instructors, four objectives of comparable difficulty were identified for the case study and placed in pairs on opposite sides of the AST. This arrangement allowed for parallel self-guided practice sessions, where two pairs of trainees could work on distant objectives to avoid interferences. Additionally, an unused objective from the training phase, located on the opposite side of the table, was always available for the exam of the single trainee.

Finally, to enable an objective evaluation component in the practical exam, which is normally absent in the traditional course, it was agreed with the AIB Corps instructors to input the interventions proposed by the trainees during the exam into a desktop version of the XRTS (modified for the purpose, and designed to work via mouse and keyboard, without requiring the AST or the tablet). This allowed for simulating their effects and verifying whether the DP has been successfully managed. This outcome was used as an additional evaluation item, alongside the standard subjective assessment performed by the instructor. Again, for logistical reasons, the evaluation involves two trainees at a time, each working on distinct objectives that differ from those used in the training, with each trainee being supervised by a separate instructor.

It should be noted that, despite the presence of two instructors, the evaluations for each trainee are determined collectively by both the instructors. To avoid bias, the instructors, who were different from those managing the training sessions, did not know whether a trainee belonged to the AST group or the XST group. Finally, all the trainees in the experiment had the same instructors for the training and the exam.

D. Procedure

The overall experimental procedure for both the evaluated conditions is described below, and depicted in Figure 4.

1) Preparation: At the beginning of the course, the participants were introduced to the experiment. Additionally, a demographic survey was conducted to collect personal data (gender and age) and record the participants' previous experience with the sand table and XR applications. As also confirmed by the

[4]Videos of the training activities: http://tiny.cc/bgkazz

instructors, none of the trainees had prior experience in using the AST, nor in the operational procedures required during the simulated fire management phase of the exam, making a knowledge pre-test impracticable. This situation was expected, as the course was the first in the entire training program to introduce the AST and its functionalities. Afterward, both the groups participated in the same theoretical lessons concerning the forest firefighting role and the use of the AST.

2) Self-directed Practice Session on the AST: For the self-directed practice, the participants from both the groups were subdivided in pairs and experienced a 30-minute practice session on the AST under the supervision of a specific instructor.

For the AST group, the practice session was conducted on the sole AST. The traditional approach was employed, allowing the trainees to discuss their maneuvers, interact with the AST by placing flags and vectors (Fig. 3d), and ask the instructor for feedback on their maneuvers.

For the XST group, the practice session involved using the XRTS with a provided tablet. These trainees used the MR application (Fig. 3e) to align the simulation with the AST, navigate fire states, visualize objectives in 3D, inspect cell information, and draft maneuvers. In doing so, they took turns using the MR device and collaborated throughout the session. This organization ensured a balanced exposure to technology, as each trainee was required to use the tablet for half of the allocated time. Like in the traditional training, a different instructor per pair of trainees was present to provide support on the use of the MR application. Due to logistical limitations, only two pairs of trainees were allowed to practice simultaneously, each working on a different objective.

3) Measures: A set of subjective and objective metrics were collected to compare the traditional AST-based training (AST group) with the devised XRTS-based one (XST group). Subjective metrics consisted of:

- The Simulation Task Load Index (*SIM-TLX*) questionnaire [40], aimed at assessing the cognitive load associated with the training experience. The questionnaire assesses the workload during a training procedure by asking to rate the level of experienced demand on nine bipolar 5-to-100 scales (Mental demand, Physical demand, Temporal demand, Frustration, Task complexity, Situational stress, Distraction, Perceptual strain, and Task control). The weighted version of this questionnaire was used to further emphasize which scale had the greatest impact on the training.
- The *AttrakDiff* User Experience (UX) questionnaire [41], which is composed of 28 pairs of opposing terms to be evaluated on a 1-to-7 scale (with 1 corresponding to the positive term of the pair and 7 to the negative one). The 28 pairs can be grouped into four categories (Pragmatic Quality, Hedonic Quality - Stimulation, Hedonic Quality - Identification, and Attractiveness) designed to collect qualitative feedback on the learning experience.
- The System Usability Scale (*SUS*) questionnaire [42], designed to estimate the overall usability of a system. In this case, it was used to estimate the usability of the considered learning tool (the AST or the XRTS).

Regarding the objective measures, the metric that was chosen to evaluate the two conditions is the score from the evaluation of the practical exam, which tests the knowledge acquired through the experience on the AST along with the time taken by the trainee to provide the solution for the exam, referred to as the examination time. The parameters considered for the evaluation encompassed the safety, effectiveness, and efficiency of the interventions within the responsibility of team leaders (i.e., involving pumps and fire trucks) scheduled by the trainee, together with general considerations on the timing, safety, and correctness of the whole experience. Since protecting the assigned objective required a preventive aircraft attack to create the necessary operative conditions, the instructor signaled an error whenever a trainee failed to request this intervention in their maneuver. The solution was then provided, and the trainee was asked to continue the examination under the correct conditions. Finally, as mentioned earlier, the outcome of the maneuvers, evaluated by inputting the relative interventions into the desktop version of the XRTS (Fig. 3f) and running a simulation, was also used as an evaluation metric. The full questionnaire and the evaluation sheet used during the practical exam are available online[5].

VI. Results

Results collected using the metrics presented in the previous section were used to compare the performance of the AST and XST groups. To analyze the statistical significance of the results, the Shapiro-Wilk test was performed to check the normality of the data. Since data proved to be characterized by non-normal distributions, pairwise comparisons were analyzed by using the non-parametric Mann-Whitney U test for two independent samples.

A. Objective Results

Regarding the objective evaluation, the exam items and the corresponding scores obtained by the two groups are reported in Table II. Each trainee was closely observed by an instructor, and a point was assigned to each correct item according to his or her behavior during the exam. The exam encompassed two parts: a first part evaluating the team-leader-specific interventions (items 1–4) and a second part evaluating overall fire management (items 5–7); for each part, the score corresponds to the sum of each correct item and is expressed as a percentage.

Starting from the first part, it is possible to observe that the XST group performed significantly better than the AST group; specifically, the trainees in the XST group scheduled direct attacks that were significantly more effective (item 1, $p = .038$) and significantly safer (item 3, $p = .038$). No differences were found for what it concerns the efficiency of the interventions (item 2, $p = .085$), since the majority of the trainees were not able to optimize resource usage regardless of the group; similar considerations apply to the ability to choose the appropriate intervention (item 4, $p = .079$).

Considering the second part, it is again possible to observe that the XST group performed significantly better than the

[5]Questionnaire and exam evaluation sheet: http://tiny.cc/0gkazz

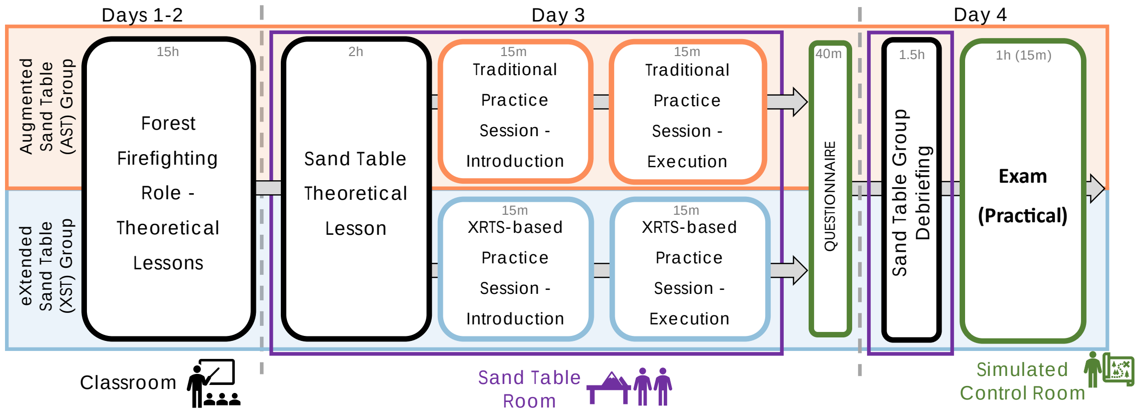


Fig. 4. Experimental procedure followed by the AST group (upper part) and XST group (lower part).

AST group. The trainees in both the groups understood that the initial conditions required a preventive intervention (i.e., an aircraft attack), in the first step of the exam (item 5, $p = .163$). The majority of the trainees were also able to successfully protected the assigned objective (item 7, $p = .163$). However, looking at the correctness of the scheduled interventions, the XST group performed significantly better than the AST group (item 6, $p = .018$), being able to request more frequently a feasible maneuver of two interventions (i.e., the initial intervention choice enabled the execution of the second one).

The partial scores for the two parts were used to calculate the total score of the exam, which was then expressed on a 1-to-100 scale. Overall, the XST group performed significantly better than the AST group. Moreover, looking at the total examination time, it is possible to observe that the trainees in the XST group managed to achieve these results in significantly less time ($\overline{x} = 9'48''$, $s = 2'43''$ for the AST group, $\overline{x} = 5'37''$, $s = 1'21''$ for the XST group, with $p \leq .001$).

TABLE II
RESULTS FOR THE PRACTICAL EXAM. SIGNIFICANT DIFFERENCES ($p < .050$) AND BEST RESULTS ARE HIGHLIGHTED IN BOLD.

Team leader-specific interventions	**AST**	**XST**
1. Carried out the interventions effectively	75.00%	100%
2. Carried out the interventions efficiently	31.25%	62.25%
3. Carried out the interventions safely	75.00%	100%
4. Carried out appropriate interventions	81.00%	100%
Score	$\mu = 65.62\%$ $\delta = 29.15$	$\mu = \mathbf{90.62}\%$ $\delta = 12.10$
p-value	$p = \mathbf{.003}$	
Overall fire management	**AST**	**XST**
5. Requested the preventive intervention (i.e., aircraft attack)	87.50%	100%
6. Requested all the interventions correctly	68.75%	100%
7. Protect the assigned objective	87.50%	100%
Score	$\mu = 81.25\%$ $\delta = 33.27$	$\mu = \mathbf{100}\%$ $\delta = 0$
p-value	$p = \mathbf{.018}$	
Total score	**AST**	**XST**
Score	$\mu = 72.32\%$ $\delta = 26.95$	$\mu = \mathbf{94.64}\%$ $\delta = 6.92$
p-value	$p = \mathbf{.003}$	

B. Subjective Results

Regarding the subjective evaluation, results for *SIM-TLX* are shown in Fig. 5. Statistically significant differences were found for three sub-scales (Fig. 5a) and the overall cognitive load, computed as a weighted average of all the sub-scales (Fig. 5b). Specifically, the XST group perceived a dramatically lower temporal demand than the AST group ($\overline{x} = 291.87$, $s = 179.06$ for the AST group, and $\overline{x} = 89.06$, $s = 98.39$ for the XST group, with $p = .001$). Moreover, the XST group experienced considerably lower task complexity compared to the AST group ($\overline{x} = 248.75$, $s = 162.42$ for the AST group, and $\overline{x} = 123.12$, $s = 74.53$ for the XST group, with $p = .020$). It is also possible to note that the XST group was subject to lower stress than the AST group ($\overline{x} = 216.87$, $s = 179.11$ for the AST group, and $\overline{x} = 100.93$, $s = 97.55$ for the XST group, with $p = .025$). Finally, the mean cognitive load was much lower for the XST group ($\overline{x} = 179.48$, $s = 61.57$ for the AST group, and $\overline{x} = 107.60$, $s = 42.99$ for the XST group, with $p = .004$).

For what it concerns the *AttrakDiff* questionnaire, the XST group described the UX with the training tool as more human and less technical than the AST group ($\overline{x} = 5.87$, $s = 1.21$ for the AST group, $\overline{x} = 4.12$, $s = 1.76$ for the XST group, with $p = .007$), probably since the use of the XRTS reduces the time spent on cumbersome tasks like, e.g., vector calculations, allowing more time for discussion between the trainees. No other significant differences were found.

Lastly, regarding the *SUS*, the score assigned to the overall usability of the adopted training tool was $\overline{x} = 69.69$, $s = 14.60$ and $\overline{x} = 70.00$, $s = 10.00$, respectively for the AST and XST groups, with no significant differences between them. According to the categorization proposed by Bangor et al. [43], both the evaluations correspond to the C grade and are associated with the "Good" usability class.

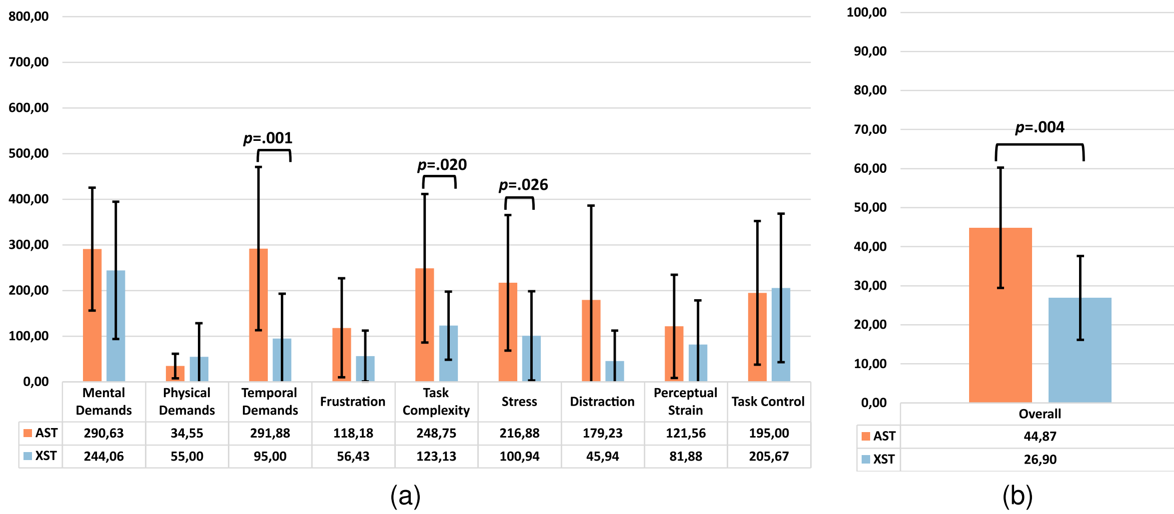


Fig. 5. Results regarding cognitive load obtained using the *SIM-TLX* questionnaire [40] (the lower the better): (a) scores for the individual sub-scales and (b) overall score obtained computing a weighted average of all the sub-scales. The sub-scales can range up to a maximum of 800, whereas the overall score is normalized in the 0–100 range. Brackets denote statistically significant differences ($p < .05$), whereas bars represent standard deviations.

VII. Discussion and Conclusion

This paper studied how to improve the training of AIB IC figures in tactical and strategic management using immersive technologies and, specifically, an XRTS. The paper contributes to advancing the research in this field by reporting the results of an investigation based on two hypotheses:

- **H1:** the use of an XRTS fosters decision-making skills on large-scale firefighting scenarios, thus improving the assimilation of the procedural details of the considered tasks and enabling the trainees to correctly manage the AIB Corps maneuvers and schedule interventions better than with the traditional AST-based training.
- **H2:** the use of an XRTS has a lower impact in terms of overall cognitive load perceived by the targeted figures during the training experience.

The two hypotheses were investigated through the analysis of post-experience measures collected through a user study.

Regarding H1, looking at the overall results it appears that the trainees utilizing the XRTS (XST group) generally outperformed those who underwent the traditional training (AST group), achieving full scores in all the items of the exam evaluation except for interventions efficiency. Similar performance trends were observed for the AST group but only with good scores across the various items. Lower ratings in both the groups regarding efficiency were consistently linked to excessive scheduling of ground resources. This finding seems reasonable, given that firefighting resource management is not a primary focus of the team leaders training and that restrictions on ground resources are typically lax. Importantly, the effectiveness of the maneuvers was not impacted. It is worth noting that, in this context, the addition of a quantitative pre-evaluation could further confirm the benefits associated with the use of the XRTS.

Considering other evaluated aspects, the XRTS appears to provide additional advantages over traditional training. Specifically, while most of the trainees in the AST group successfully protected the assigned objective, requested the correct preventive aircraft attack, and chose appropriate interventions, they generally struggled with scheduling the correct combination of interventions (i.e., requesting an aircraft attack, then a direct attack). As a result, the performance of the ATS group was observed to be notably lower than that of the XST group.

Moreover, it was frequently observed that trainees in the AST group did not request the preemptive aircraft attack before proceeding with other maneuvers, which may be related to their underestimation of the wildfire intensity. In other cases, it was observed that after the initial aircraft pass, trainees in the AST group scheduled a second aircraft attack instead of a direct one, which may be related to an overestimation of the fire intensity. This choice, although successful, is not standard practice in real-life operations. These errors led to the choice of an incorrect maneuver.

Moving to the detailed analysis of scheduled interventions, it is evident that most of the trainees in the AST group successfully requested preemptive aircraft attacks, with no significant differences with respect to trainees in the XST group. Thus, the notably lower score was reasonably due to challenges with subsequent interventions, which are also under the responsibility of team leaders. Specifically, training with the XRTS seemed to improve effectiveness and safety in interventions involving pumps and fire trucks. At the IC level, scheduling such interventions is not immediate and depends on fire intensity and terrain logistics around the objective, with variable interventions parameters.

Overall, the trainees who used the XRTS achieved significantly better exam results compared to those who underwent the traditional training. The use of the XRTS, particularly the MR application's direct interaction with the simulation and the timely feedback from the simulator, was found to be beneficial in enhancing critical thinking and preparing the trainees for effective maneuvers. This outcome aligns with previous research on MR applications in different and simpler learning scenarios [44], [45], [46], and supports the hypothesis. These advantages in critical thinking and preparation appeared to have an immediate positive impact on exam performance (as seen in [47]), with the XST group needing half the time of the AST group to complete the task. In this context, it is worth noting that the automatic variable calculation for decision-making in the XRTS may have simplified the implementation of maneuvers compared to the traditional AST, regardless of the use of the simulator and XR. However, this effect is expected to be marginal due to the study's design, as the exam was conducted without the use of the XRTS and both the groups followed the same procedure.

For what it concerns H2, looking at the obtained results, it is possible to observe that the XST group apparently showed less susceptibility to task complexity compared to the AST group, likely because, in the former case, the trainees used the simplified and intuitive simulation interface provided by the XRTS. They were also less affected by time demand and reported lower stress levels during training, a result that can probably be attributed to XRTS features that seemingly facilitated information retrieval and clear observation of maneuvers outcomes. Overall, the trainees who used the XRTS reported experiencing much lower cognitive load compared to the trainees who underwent traditional training, supporting the hypothesis. In a conflicting research scenario where the impact of XR on cognitive load is largely debated [48], results reported in this paper seem to contradict those studies that highlight the negative impact of immersive technologies [49], [50], suggesting that XR-based experiences can be associated with a reduced load, as reported, e.g., in the review by Buchner et al. [51]. Again, the automatic variable calculation for decision-making in the XRTS may have positively biased the XST group by simplifying the practice session. However, this effect should have been limited, as instructors in the AST training provided similar feedback to prevent major errors, offering a comparable support.

These considerations are further supported by the results regarding the overall UX, which revealed positive judgments from both the groups in terms of Pragmatic Quality, Hedonic Quality, and Attractiveness. The XRTS-based experience, however, offered a simplified control interface that seemingly reduced perceived task complexity and appeared to address the limitations of the AST. The trainees of the XST group found the XRTS more user-friendly, allowing them to engage in the training without dealing with technical details or complex actions, and probably leading to a smoother and more efficient training experience than with the AST.

To conclude, based on the above considerations, the findings of this work can be summarized as follows: the XRTS apparently provided the trainees with a better overall view of the considered firefighting scenario and the necessary means to select the most appropriate strategic approach, thus potentially improving the decision-making process and enabling them to schedule more accurate maneuvers (i.e., combinations of interventions), which supports H1. Moreover, the use of the XRTS was observed to have a positive impact on overall cognitive load, perceived stress, perceived task complexity, and sense of time passing, making the experience appear more efficient and less time-consuming. This outcome supports H2.

It is important to acknowledge that the conclusions of this study are subject to the limited sample size, which was constrained by the structure of the training course. Specifically, the relatively small number of participants (n=32) resulted from the decision to limit the experiment to a single course edition. This approach was adopted to minimize potential confounding factors, since both the AST configuration (i.e., the case study) and instructor assignments vary across different editions, potentially introducing inconsistencies. By conducting the study within a single edition, we ensured a controlled experimental environment, thereby enhancing the reliability of the findings. Nevertheless, despite this limitation, the study offers valuable insights, and contributes to the ongoing exploration of XR and interactive simulations in training contexts, highlighting the potential advantages of these technologies in educational settings compared to traditional methods.

A second limitation concerns the automatic variable calculation for decision-making in the XRTS, which may have simplified maneuvers implementation compared to the traditional AST, regardless of simulator or XR use. This feature might have advantaged the XST group, particularly reducing cognitive load and decision-making complexity. However, in the AST training, instructors provide similar feedback to prevent major errors, thus offering a comparable support. Moreover, since the XRTS was not used during the exam, both the groups had equal access to resources. Hence, the automatic calculation feature likely had little to no effect on exam outcomes and may have even hindered performance, as decision-making had to be conducted independently. A detailed breakdown analysis of the XRTS [52] could further clarify how each component influences overall outcomes, providing deeper insights into its role in decision-making performance.

The absence of a quantitative pre-evaluation at the beginning of the experiment could be seen as a limitation too; however, this decision was based on the fact that all the participants were novice trainees with no prior AST experience, as confirmed by the course instructors. Nevertheless, a pre-evaluation could be particularly beneficial in refresher courses, where trainees have prior experience with the AST but no exposure to the XRTS; by incorporating it to establish a baseline for assessing learning gains, future studies in this context could provide deeper insights into the results and enable a more structured evaluation of training effectiveness.

Another area of exploration for future studies concerns the broader application of the XRTS in IC contexts. In this study, the XRTS was assessed only for team leaders, who focused on tactical aspects using a limited subset of interventions over a restricted simulation period (two steps of one hour each). While this limitation is unlikely to have affected the results, further studies should deploy and evaluate the XRTS in formal courses designed for other IC roles, extending the analysis to strategic decision-making and incorporating the full range of supported interventions.

A potential improvement of the XRTS lies in the management of interventions by the fire simulator. Currently, each intervention is considered fully executed once scheduled, without accounting for human errors or unforeseen events such as new ignition points or changing external factors (e.g., weather conditions). This approach aligns with traditional methods, as instructors followed the same procedure during conventional practice sessions. To enhance functionality, an extended XRTS version could estimate the likelihood of successful intervention execution and allow instructors to apply scenario modifications (i.e., injections) to introduce unpredictability and assess trainees' responses to unexpected situations.

Finally, the choice to use tablets aimed to mitigate the technology barrier, as they are familiar devices. However, this benefit was not observed in practice, since using tablets as MR devices deviated from their traditional use and diverted trainees' attention from the AST, with the potential loss of key perceptual cues. While trainees were encouraged to share the tablet to foster collaboration (and this goal was achieved), relying on a single device may have led to a more isolating experience. Future studies could explore an XRTS where each trainee a tablet in order to assess its impact on collaboration. This approach presents challenges, including the need for simulation synchronization and the potential distraction from the AST. While this setup may be suitable for individual training, it might not align with the pair-based format required by the course. Alternatively, HMDs could resolve the single-device issue while refocusing attention on the AST.

Future work could also explore replacing handheld devices with HMDs to enhance MR content visualization and interaction with the XRTS, considering both see-through (e.g., Microsoft HoloLens 2) and video-see-through devices (e.g., Meta Quest Pro and HTC Vive XR Elite). To this aim, the system will need modifications to support a multi-user experience, letting multiple trainees to interact with the same simulation from different viewpoints and devices. This transition will also introduce new usage modes, enabling users with different roles (e.g., team leaders, CoAIBs, analysts) to collaborate simultaneously, either co-located or remotely, within the same case study and simulation. Expanding interaction across sub-zones and command levels could enable new training approaches that surpass the performance of current methods.

Lastly, it may be interesting to assess the performance of the devised XRTS in terms of learning retention, contrasting results against the traditional course. This would request to involve again the participants of the experimental activity, e.g., during scheduled refresher courses, and request them to reapply what they learned and recall from prior experiences.

Acknowledgments

The authors want to thank Giovanni Folcarelli and Luca Bellardi Gioli for their contribution to the development of the XRTS and to acknowledge the support provided by AIB Corps, in particular by Sergio Pirone and Augusto Cotterchio, in the design and validation phases. Finally, they would like to thank Alessandro Visconti and Roberta Macaluso for their support during the experiments.

References

[1] J. C. Uhl, G. Regal, H. Schrom-Feiertag, M. Murtinger, and M. Tscheligi, "XR for first responders: Concepts, challenges and future potential of immersive training," in Virtual Reality and Mixed Reality, G. Zachmann, K. Walczak, O. A. Niamut, K. Johnsen, W. Stuerzlinger, M. Alcañiz-Raya, G. Welch, and P. Bourdot, Eds., 2023, pp. 192–200.

[2] K. Lamb, M. Farrow, C. Olymbios, D. Launder, and I. Greatbatch, "Systematic incident command training and organisational competence," Int. J. of Emergency Services, vol. 10, no. 2, pp. 222–234, 2021.

[3] J. Jensen and S. Thompson, "The incident command system: A literature review," Disasters, vol. 40, no. 1, pp. 158–182, 2016.

[4] K. Jane Lamb, J. Davies, R. Bowley, and J.-P. Williams, "Incident command training: The introspect model," Int. J. of Emergency Services, vol. 3, no. 2, pp. 131–143, 2014.

[5] R. Chang, "A literature review and analysis of the incident command system," Int. J. of Emergency Management, vol. 13, p. 50, 01 2017.

[6] J. Haskins, B. Zhu, S. Gainer, W. Huse, S. Eadara, B. Boyd, C. Laird, J. Farantatos, and J. Jerald, "Exploring VR training for first responders," in Proc of. IEEE Conf. on Virtual Reality and 3D User Interfaces Abstracts and Workshops, 2020, pp. 57–62.

[7] F. Marincioni, "Information technologies and the sharing of disaster knowledge: The critical role of professional culture," Disasters, vol. 31, no. 4, pp. 459–476, 2007.

[8] Z. Zhou, Z. Bian, and Z. Zhuo, "MR sand table: Mixing real-time video streaming in physical models," in Proc. of IEEE Virtual Reality. IEEE, 2017, pp. 239–240.

[9] M. Kirschenbaum, "Granular worlds: Situating the sand table in media history," Critical Inquiry, vol. 50, no. 1, pp. 137–163, 2023.

[10] T. L. Woods, S. Reed, S. Hsi, J. A. Woods, and M. R. Woods, "Pilot study using the augmented reality sandbox to teach topographic maps and surficial processes in introductory geology labs," J. of Geoscience Education, vol. 64, no. 3, pp. 199–214, 2016.

[11] Y. Qingshu, W. Ruonan, P. Zhigeng, X. Shuchang, G. Jiali, and L. Tianren, "A survey on human-computer interaction in spatial augmented reality," J. of Computer-Aided Design & Computer Graphics, vol. 33, no. 3, pp. 321–332, 2021.

[12] J. M. F. Rodrigues, P. J. S. Cardoso, J. Monteiro, and C. M. Q. Ramos, Smart Systems Design, Applications, and Challenges. Hershey, PA, USA: IGI Global, 2020, pp. 1–459.

[13] D. Mathiesen, T. Myers, I. Atkinson, and J. VanDerWal, "Terrainvis: Collaborative, interactive, visualisation tool for engaging residents in disaster preparedness," in Proc. of Cooperative Design, Visualization, and Engineering: 13th Int. Conf. Springer, 2016, pp. 110–117.

[14] N. Pellas, A. Dengel, and A. Christopoulos, "A scoping review of immersive virtual reality in stem education," IEEE Trans. on Learning Technologies, vol. 13, no. 4, pp. 748–761, 2020.

[15] Y. Cao, G.-W. Ng, and S.-S. Ye, "Design and evaluation for immersive virtual reality learning environment: A systematic literature review," Sustainability, vol. 15, no. 3, 2023.

[16] D. Paes, Z. Feng, S. Mander, S. Datoussaid, T. Descamps, A. Rahouti, R. Lovreglio et al., "Video see-through augmented reality fire safety training: A comparison with virtual reality and video training," Safety Science, vol. 184, p. 106714, 2025.

[17] G. Koutitas, S. Smith, and G. Lawrence, "Performance evaluation of AR/VR training technologies for EMS first responders," Virtual Reality, vol. 25, no. 1, pp. 83–94, 2021.

[18] D. Calandra, F. G. Pratticò, M. Migliorini, V. Verda, and F. Lamberti, "A multi-role, multi-user, multi-technology virtual reality-based road tunnel fire simulator for training purposes," in Proc. of 16th Int. Conf. on Computer Graphics Theory and Applications, 2021, pp. 96–105.

[19] M. P. Simanjuntak, J. Hutahaean, N. Marpaung, and D. Ramadhani, "Effectiveness of problem-based learning combined with computer simulation on students' problem-solving and creative thinking skills." Int. J. of Instruction, vol. 14, no. 3, pp. 519–534, 2021.

[20] F. Corelli, E. Battegazzorre, F. Strada, A. Bottino, and G. P. Cimellaro, "Assessing the usability of different virtual reality systems for firefighter training," in Proc. of 15th Int. Joint Conf. on Computer Vision, Imaging and Computer Graphics Theory and Applications, 2020, pp. 146–153.

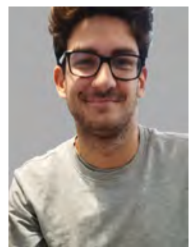
[21] H. Engelbrecht, R. W. Lindeman, and S. Hoermann, "A swot analysis of the field of virtual reality for firefighter training," Frontiers in Robotics and AI, vol. 6, 2019.

[22] M. Castellnou et al., "Empowering strategic decision-making for wildfire management: Avoiding the fear trap and creating a resilient landscape," Fire ecology, vol. 15, pp. 1–17, 2019.

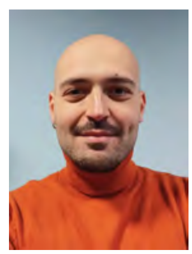
[23] G. Klein, R. Calderwood, and A. Clinton-Cirocco, "Rapid decision making on the fire ground: The original study plus a postscript," J. of cognitive engineering and decision making, vol. 4, no. 3, pp. 186–209, 2010.

[24] D. Calandra, F. De Lorenzis, A. Cannavò, and F. Lamberti, "Immersive virtual reality and passive haptic interfaces to improve procedural

learning in a formal training course for first responders," Virtual Reality, vol. 27, no. 2, pp. 985–1012, 2023.
[25] C. Pais, J. Carrasco, D. L. Martell, A. Weintraub, and D. L. Woodruff, "Cell2fire: A cell based forest fire growth model," arXiv preprint arXiv:1905.09317, 2019.
[26] O. Zechner, D. García Guirao, H. Schrom-Feiertag, G. Regal, J. C. Uhl, L. Gyllencreutz, D. Sjöberg, and M. Tscheligi, "Nextgen training for medical first responders: Advancing mass-casualty incident preparedness through mixed reality technology," Multimodal Technologies and Interaction, vol. 7, no. 12, 2023.
[27] H. Chen, L. Hou, G. K. Zhang, and S. Moon, "Development of BIM, IoT and AR/VR technologies for fire safety and upskilling," Automation in Construction, vol. 125, p. 103631, 2021.
[28] C. R. Amburn, N. L. Vey, M. W. Boyce, and J. R. Mize, "The Augmented REality Sandtable (ARES)," US Army Research Laboratory, Tech. Rep. ARL-SR-0340, 2015.
[29] J. Abich, M. Eudy, J. Murphy, C. Garneau, Y. Raby, and C. Amburn, "Use of the Augmented REality Sandtable (ARES) to enhance Army CBRN training," in Proc of HCI Int. Springer, 2018, pp. 223–230.
[30] M. Korkiakoski, K. Chakal, R. Fernando, and P. Kostakos, "ARTS: Augmented reality tactical sandbox for improved situational awareness in a multi-user scenario," Research Square, pp. 1–49, 2023.
[31] C.-C. Mao and C.-H. Chen, "Augmented reality of 3D content application in common operational picture training system for army," Int. J. of Human–Computer Interaction, vol. 37, no. 20, pp. 1899–1915, 2021.
[32] S. Guerin and F. Carrera, "Sand on fire: An interactive tangible 3D platform for the modeling and management of wildfires," WIT Trans. on Ecology and the Environment, vol. 137, pp. 57–68, 2010.
[33] P. L. Andrews, "BehavePlus fire modeling system: Past, present, and future," in Proc. of 7th Symp. on fire and forest meteorology. American Meteorological Society Boston MA:, 2007, pp. 23–25.
[34] M. Forestry Canada Fire Danger Group, "Development and structure of the canadian forest fire behavior prediction system," Forestry Canada, Headquarters, Fire Danger Group and Science and Sustainable Development Directorate,(Ottawa), 1992.
[35] C. Tymstra et al., "Development and structure of Prometheus: The Canadian wildland fire growth simulation model," Natural Resources Canada, Canadian Forest Service, Northern Forestry Centre, Information Report NOR-X-417, 2010.
[36] D. Campbell, The Campbell Prediction System: A Wild Land Fire Prediction System and Language. Self-Published, 1993.
[37] B. B. Baker and E. T. Copson, The mathematical theory of Huygens' principle. American Mathematical Soc., 2003, vol. 329.
[38] Kronnect, "Terrain Grid System," https://forum.unity.com/threads/terrain-grid-system-powerful-and-flexible-voronoi-boxed-and-hexagonal-grid.366248/, 2015, accessed: 2024-07-01.
[39] Ultra Real, "Ultra Real Fire Effects Volume 1 : Static Fire Props," https://forum.unity.com/threads/ultra-real-fire-effects-volume-1-static-fire-props-and-more.291539/, 2015, accessed: 2024-07-01.
[40] D. Harris, M. Wilson, and S. Vine, "Development and validation of a simulation workload measure: The simulation task load index (SIM-TLX)," Virtual Reality, vol. 24, no. 4, p. 557 – 566, 2020.
[41] M. Hassenzahl, F. Koller, and M. Burmester, "Der user experience (UX) auf der spur: Zum einsatz von www.attrakdiff.de," in Tagungsband UP08, H. Brau, S. Diefenbach, M. Hassenzahl, F. Koller, M. Peissner, and K. Röse, Eds. Stuttgart: Fraunhofer Verlag, 2008, pp. 78–82.
[42] J. Brooke, "SUS-A quick and dirty usability scale." Usability evaluation in industry. CRC Press, June 1996.
[43] A. Bangor, P. Kortum, and J. Miller, "Determining what individual SUS scores mean: Adding an adjective rating scale," J. of Usability Studies, vol. 4, no. 3, pp. 114–123, 2009.
[44] H. Faridi, N. Tuli, A. Mantri, G. Singh, and S. Gargrish, "A framework utilizing augmented reality to improve critical thinking ability and learning gain of the students in physics," Computer Applications in Engineering Education, vol. 29, no. 1, pp. 258–273, 2021.
[45] M. E. C. Santos, A. Chen, T. Taketomi, G. Yamamoto, J. Miyazaki, and H. Kato, "Augmented reality learning experiences: Survey of prototype design and evaluation," IEEE Trans. on Learning Technologies, vol. 7, no. 1, pp. 38–56, 2014.
[46] N. F. Saidin, N. D. Abd Halim, N. Yahaya, and N. N. Zulkifli, "Enhancing students' critical thinking and visualisation skills through mobile augmented reality," Knowledge Management & E-Learning, vol. 16, no. 1, pp. 1–41, 2024.
[47] F. De Lorenzis, F. G. Pratticò, M. Repetto, E. Pons, and F. Lamberti, "Immersive virtual reality for procedural training: Comparing traditional and learning by teaching approaches," Computers in Industry, vol. 144, p. 103785, 2023.
[48] M. Akçayır and G. Akçayır, "Advantages and challenges associated with augmented reality for education: A systematic review of the literature," Educational Research Review, vol. 20, pp. 1–11, 2017.
[49] G. Makransky and G. B. Petersen, "Investigating the process of learning with desktop virtual reality: A structural equation modeling approach," Computers & Education, vol. 134, pp. 15–30, 2019.
[50] G. B. Petersen, A. Mottelson, and G. Makransky, "Pedagogical agents in educational VR: An in the wild study," in Proc. of CHI Conf. on Human Factors in Computing Systems, 2021, pp. 1–12.
[51] J. Buchner, K. Buntins, and M. Kerres, "The impact of augmented reality on cognitive load and performance: A systematic review," J. of Computer Assisted Learning, vol. 38, no. 1, pp. 285–303, 2022.
[52] F. G. Pratticò, A. Cannavò, D. Calandra, and F. Lamberti, "A breakdown study of a mockup-based consumer haptic setup for virtual reality," IEEE Consumer Electronics Magazine, vol. 12, no. 5, pp. 76–90, 2023.



Lorenzo Valente is a PhD student at the Dipartimento di Automatica e Informatica of the Politecnico di Torino, Italy, doing research on using XR to create experiences in the domains of digital reconstruction, simulation, and interaction techniques. Contact him at lorenzo.valente@polito.it.



Federico De Lorenzis received the B.Sc., M.Sc. and Ph.D. degrees in computer engineering from Politecnico di Torino, Italy, in 2016, 2019 and 2024, respectively. He is currently a research fellow at the Dipartimento di Automatica e Informatica of Politecnico di Torino. His research activities focus on the use of VR and AR for the training of users in both formal and non-formal contexts. Contact him at federico.delorenzis@polito.it.

Davide Calandra received the B.Sc., M.Sc., and Ph.D. degrees in computer engineering from Politecnico di Torino, Italy, in 2014, 2017, and 2023. He is a junior assistant professor at the Dipartimento di Automatica e Informatica of Politecnico di Torino, with interests in interactive graphics applications, XR and human-machine interaction. He is a member of IEEE and IEEE Consumer Technology Society. Contact him at davide.calandra@polito.it.

Fabrizio Lamberti received the M.Sc. and Ph.D. degrees in computer engineering from Politecnico di Torino, Italy, in 2000 and 2005, respectively. He is a full professor at the Dipartimento di Automatica e Informatica of Politecnico di Torino, where he leads the VR@POLITO lab. His research interests include computer graphics, human-machine interaction, and computational intelligence. He serves as an Associate Editor for several journals, including IEEE Transactions on Computers, IEEE Transactions on Consumer Electronics, and International Journal of Human-Computer Studies. He is a Member of the Board of Governors of IEEE Consumer Technology Society serving as Vice President of Technical Activities.