

Preliminary Evaluation of an Asymmetric Mixed Reality Firefighter Training System

Katarina Mišura^{1,*}, Mirko Sužnjević¹, Lea Skorin-Kapov¹ and Maja Matijašević¹

¹University of Zagreb Faculty of Electrical Engineering and Computing, Unska 3, Zagreb, 10000, Croatia

Abstract

Firefighter training benefits from realistic and repeatable exercises, while conventional live-fire training can impose safety, logistical, and resource constraints. This paper presents an evaluation of a mixed reality (MR) training system designed to complement physical firefighter training. In an MR exercise, firefighters physically navigate a virtual house registered within the training space, locate simulated fires, and extinguish them using a handheld controller. The system combines an MR application for trainees with a desktop application that enables instructors to configure, initiate, and monitor training scenarios. The system was evaluated in a study involving 65 professional firefighters, including regular firefighters (as trainees) and instructors qualified to lead or supervise training exercises. The evaluation examined the experience from both trainee and instructor perspectives and included measures of usability, technology acceptance, simulator sickness, technical limitations, instructor information requirements, and perceived educational value. This paper reports the study design and preliminary findings, with a focus on the perceived relevance of the exercise for professional firefighter training.

Keywords

mixed reality, firefighter training, extended Reality, training simulation, immersive training

1. Introduction and Related Work

Firefighting involves hazardous and rapidly changing environments in which effective training is essential for safe operation [1]. However, real-world firefighter training can involve safety, logistical, and resource constraints. XR technologies have therefore been explored as a means of augmenting or complementing conventional training by enabling controlled, repeatable scenarios without exposing trainees to all the hazards of real incidents [2].

Research on XR-based firefighter training has predominantly focused on virtual reality (VR). Existing reviews highlight promising applications in different training scenarios, while also emphasizing human factors, user experience, and the need for more evidence regarding training effectiveness [3, 4, 5]. Recent reviews have also emphasized the growing importance of end-user experience in VR firefighter training, including subjective and physiological responses [5]. A recent empirical study with professional users has also reported high acceptance of VR-based firefighter training, reinforcing the relevance of considering user acceptance alongside training performance [6]. Augmented and mixed reality (AR/MR) offers a different approach by integrating virtual elements into the physical training environment. Previous work has shown how AR can augment existing firefighter training scenarios [7], while other studies have examined usability, user experience, and technology acceptance in AR-based firefighter training [8]. Together, these findings support the evaluation of XR training not only in terms of functionality but also from the perspective of its intended users. The asymmetric setup can also be viewed as a form of XR-mediated human–human collaboration, in which the trainee acts within the MR scenario while the instructor maintains a complementary desktop overview of the same exercise.

In our previous work, we introduced three interrelated use cases spanning firefighter MR training, high-fidelity VR location familiarization, and real-time digital twin supported field operations [9]. This paper focuses on the MR training use case, and presents the preliminary results of a user study involving

2nd International XR4CE Workshop, November 3, 2026, Valencia, Spain

*Corresponding author.

✉ katarina.misura@fer.unizg.hr (K. Mišura); mirko.suznjedic@fer.unizg.hr (M. Sužnjević); lea.skorin-kapov@fer.unizg.hr (L. Skorin-Kapov); maja.matijasevic@fer.unizg.hr (M. Matijašević)



© 2022 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

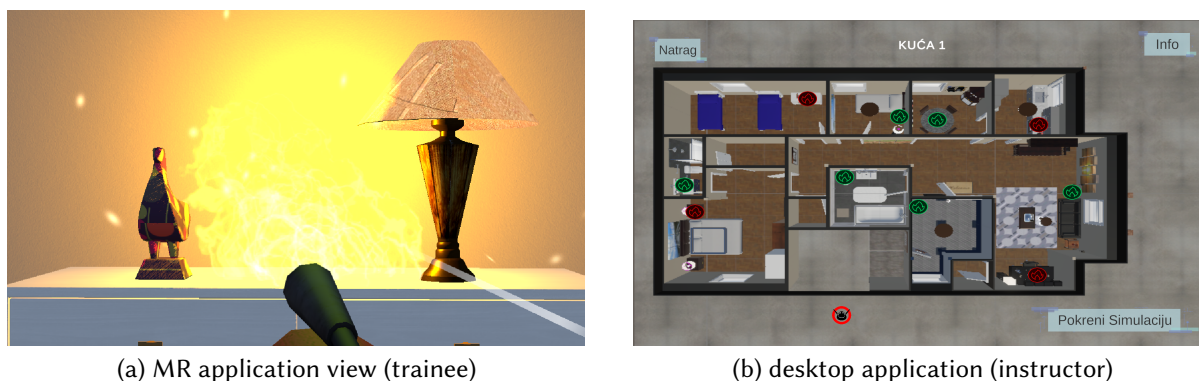


Figure 1: The asymmetric MR firefighter training system: (a) a trainee view through the MR application and (b) a desktop application used by the exercise instructor.

65 professional firefighters, focused on the evaluation of practical use, perceived training relevance, and instructor support for monitoring and assessment.

2. Methodology

The study involved evaluating an asymmetric mixed reality training system in which a firefighter trainee interacted with a spatially registered MR scenario while an instructor configured and monitored the exercise through a desktop application. This section describes the training system, the experimental task and procedure, the participants, the questionnaires, and subjective and objective measures used in the evaluation.

2.1. Training System

The training system, shown in Figure 1, consists of two coordinated applications: (a) an MR application used by the trainee to interact with the simulated environment, and (b) a desktop application used by the exercise instructor for configuring and monitoring the exercise. The MR application runs on XREAL optical see-through AR glasses and presents a virtual, fully furnished 3D house spatially registered within the physical training area. Trainees must physically move through the space, locate simulated fires, and extinguish them using a handheld controller. Upon completion of the exercise, the applications displays the number of successfully extinguished fires and the elapsed time. The desktop application allows the instructor to select and preview the training environment, activate a subset of predefined fire locations, initiate the exercise, and monitor its progress. Before the exercise begins, the desktop application verifies the MR connection and allows the instructor to confirm the trainee's starting position and readiness. During the exercise, the instructor can monitor the trainee's position, the state of each fire, the number of extinguished fires, and the elapsed time using a floor-plan view.

2.2. Experimental Task and Procedure

For the study, the training scenario was kept consistent across sessions by using the same virtual house model and four predefined fire locations. At the start of each session, the trainee put on the AR glasses, picked up the handheld controller, and moved to the designated starting area, where instructions were displayed in the MR application. The instructor selected the house model, activated the four predefined fire locations, and selected the *Start Simulation* button (in Croatian: *Pokreni Simulaciju*), as shown in Figure 1(b). A status window then indicated whether each of three conditions was met: an established MR connection, correct trainee positioning within the starting area, and trainee readiness. Once the connection and starting-position conditions were met, the instructor verbally instructed the trainee to press *Start* in the MR application to confirm readiness. The virtual house then appeared, and the

exercise began. The trainee searched the virtual house for fires. No route was prescribed, allowing participants to choose their own search strategy. They were instructed not to walk through virtual walls or furniture and were not permitted to ask the instructor for the fire locations. When a fire was found, the trainee aimed the handheld controller at it and pressed the controller to activate a simplified virtual fire extinguisher until the fire was registered as extinguished. During the exercise, the instructor monitored the trainee's position, the state of each fire, number of extinguished fires, and the elapsed time, without providing guidance on fire locations. The exercise ended when all four fires had been extinguished, but could also be terminated manually if necessary. All participants were informed that they could pause or stop the exercise at any time and were instructed to report discomfort immediately. After completion, they reviewed the number of extinguished fires and exercise duration and completed the corresponding post-study questionnaires.

2.3. Participants

A total of 65 professional firefighters participated in the user study. Participants were divided into two groups, based on their qualification to conduct training exercises. The trainee group (*P*, from the Croatian *polaznici*) consisted of 41 firefighters who were not qualified to independently lead and supervise training exercises. The instructor group (*I*) consisted of 24 firefighters who held these qualifications. The distinction between the groups reflected their roles in the study. Participants in group *P* completed the exercise as MR trainees and evaluated only the MR application. Participants in group *I* first completed and evaluated the MR exercise, with the study administrator operating the desktop application in the instructor role. They then used and evaluated the desktop application from the instructor perspective while monitoring another participant performing the MR exercise in real time; the study administrator remained present throughout the session. This order was kept fixed so that prior exposure to the desktop floor-plan view would not influence participants' navigation or search behavior during the MR exercise. Since the desktop role during exercise execution was observational and did not involve providing guidance on fire locations or modifying the scenario, having an instructor participant perform the monitoring did not alter the trainee's prescribed task. This allowed group *I* to assess both components of the system while drawing on their experience in exercise supervision, performance evaluation, safety monitoring, and debriefing.

Participants had a mean age of 36.49 years (*P*: 31.90; *I*: 44.33). Most participants (62/65; 95.4%) had previously participated in indoor fire exercises, and 61/65 (93.8%) reported participating in practical firefighter exercises at least several times per year. Participants generally reported limited prior experience with immersive technologies, with mean ratings of 2.77 for VR and 2.75 for AR on a 7-point experience scale, and 26 participants (40.0%) reported that they had never used VR or AR devices.

2.4. Subjective and Objective Measures

All study instructions and questionnaires were administered in Croatian. Before the MR exercise, all participants completed a questionnaire covering demographics, professional background, prior AR/VR and simulator experience, susceptibility to motion sickness, and affinity toward technology interaction, assessed using the Affinity for Technology Interaction (ATI) scale [10].

Following the MR exercise, the participants completed the Simulator Sickness Questionnaire (SSQ) [11] before and after exposure, an overall Quality of Experience (QoE) rating on a seven-point scale ranging from 1 (extremely poor) to 7 (excellent), the System Usability Scale (SUS) [12], the NASA Raw Task Load Index (NASA-RTLX) [13], and an adapted Technology Acceptance Model (TAM) questionnaire covering perceived usefulness, perceived ease of use, and behavioral intention [14, 15]. Additional custom items addressed technical limitations such as tracking and alignment, field of view, visibility, controller alignment, headset comfort, physical interference, and perception of real-world obstacles. The instructors additionally evaluated the desktop application using QoE, SUS, and adapted TAM measures. Custom questions addressed the information needed to monitor and assess an exercise, including trainee's position and movement, room visitation, fire status, elapsed time, trainee actions,

Table 1

Distribution of responses to the perceived training-value items by participant group. Values are percentages. (5-pt Likert scale, anchors: Strongly disagree - Strongly agree)

Statement	Group	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
Objective was clear	All (<i>n</i> = 65)	0.00%	0.00%	1.54%	9.23%	89.23%
	<i>P</i> (<i>n</i> = 41)	0.00%	0.00%	0.00%	12.20%	87.80%
	<i>I</i> (<i>n</i> = 24)	0.00%	0.00%	4.17%	4.17%	91.67%
Tasks relevant to real training	All (<i>n</i> = 65)	0.00%	0.00%	10.77%	21.54%	67.69%
	<i>P</i> (<i>n</i> = 41)	0.00%	0.00%	7.32%	24.39%	68.29%
	<i>I</i> (<i>n</i> = 24)	0.00%	0.00%	16.67%	16.67%	66.67%
Required relevant decision-making	All (<i>n</i> = 65)	0.00%	1.54%	7.69%	29.23%	61.54%
	<i>P</i> (<i>n</i> = 41)	0.00%	2.44%	4.88%	36.59%	56.10%
	<i>I</i> (<i>n</i> = 24)	0.00%	0.00%	12.50%	16.67%	70.83%
Worth developing further	All (<i>n</i> = 65)	0.00%	0.00%	1.54%	13.85%	84.62%
	<i>P</i> (<i>n</i> = 41)	0.00%	0.00%	0.00%	19.51%	80.49%
	<i>I</i> (<i>n</i> = 24)	0.00%	0.00%	4.17%	4.17%	91.67%

technical problems, and information useful during or after the exercise. Open-ended feedback was also collected.

A common four-item block assessed perceived training value on a five-point Likert scale from 1 (strongly disagree) to 5 (strongly agree), covering exercise clarity, relevance to real training, relevant decision-making, and whether the MR exercise was considered worth developing further. These items reflect perceived training value rather than actual learning effectiveness or transfer.

3. Preliminary Results and Discussion

Given the preliminary scope of this study, the results reported here are descriptive, no inferential hypothesis testing or comparison with a control condition was performed.

All 65 participants completed the planned MR exercise. No participant requested or required early termination, and all participants reported that they had experienced no symptoms that would have led them to discontinue the exercise. No session was interrupted by a loss of connection between the MR and desktop applications.

The overall experience with the MR application was rated highly, with a mean QoE score of 6.35 on a seven-point scale (seven-point scale, anchors: 1 = extremely poor - 7 = excellent). Ratings were similar between trainees ($M = 6.37$) and instructors ($M = 6.33$). Instructors also rated the desktop application highly ($M = 6.42$). While these ratings do not by themselves establish usability or training effectiveness, they provide an initial indication that both components of the asymmetric training system were positively received.

Participants reported consistently positive perceived training value across both groups (Table 1). Responses were concentrated toward the upper end of the scale, particularly for exercise clarity and whether this type of exercise was worth further development, suggesting a possible ceiling effect. The overall response patterns were similar between groups, although instructors showed somewhat more neutral responses regarding the relevance of the tasks to real training, while showing stronger endorsement of relevant decision-making and further development. As the analysis is descriptive, no inferential comparisons between groups were performed.

The instructor responses further suggest practical applicability: 22 of 24 instructors (91.7%) reported that the information currently available through the desktop application would allow them to adequately assess the performance of the MR trainee. This suggests that the asymmetric setup may support not only scenario execution but also instructor-led observation and assessment, although the specific information

requirements for monitoring and debriefing require further analysis.

The current instructor role is primarily supervisory, comprising scenario configuration, exercise initiation, monitoring, and assessment. During the exercise, the instructor maintains an overview of the trainee's position, fire status, and progress, but does not actively intervene in the scenario or guide the trainee. Future versions could support a more active collaborative role by allowing instructors to adapt scenario elements during an exercise, introduce additional events, or provide richer information for post-exercise debriefing. Such capabilities could strengthen the interaction between the trainee and the instructor and allow the training scenario to respond more dynamically to the performance of the trainee.

Overall, these preliminary findings indicate a positive reception of the MR training system and its perceived relevance to professional firefighter training. As already mentioned, actual learning outcomes and training transfer remain to be evaluated.

4. Conclusions and Ongoing Work

The initial study indicates that MR can be integrated into a physically performed firefighter training exercise in a way that is positively received by professional firefighters and perceived as relevant to training practice. The complete dataset with collected subjective and objective measures – including usability, workload, technology acceptance, simulator sickness, MR-specific technical limitations, instructor information requirements, qualitative feedback, and role-specific analysis – is currently being analyzed and will be reported in future work. Overall, the findings provide initial support for further investigating MR as a means of complementing and augmenting, rather than replacing, physical firefighter training.

Acknowledgments

Research was funded by the European Union – Next Generation EU as part of the project “Application of Immersive Technologies, Geolocation, and Digital Twins for Improving Training and Field Support in Firefighting (IGNIS)”, grant number NPOO 581-16942. The views and opinions expressed are solely those of the authors and do not necessarily reflect those of the European Union or the European Commission. Neither the European Union nor the European Commission can be held responsible for them.

Declaration on Generative AI

The authors used ChatGPT-5.4 to check spelling and paraphrase. After using this tool/service, the authors reviewed and edited the content as needed and take full responsibility for the publication's content.

References

- [1] V. Dunn, *Safety and Survival on the Fireground*, Fire Engineering Books, 2015.
- [2] J. Heirman, S. Selleri, T. De Vleeschauwer, C. Hamesse, M. Bellemans, E. Schoofs, R. Haelterman, Exploring the possibilities of extended reality in the world of firefighting, in: 2020 IEEE international conference on artificial intelligence and virtual reality (AIVR), IEEE, 2020, pp. 266–273.
- [3] P. M. Karoti, S. Adhikari, Review of fire emergency training using virtual reality, *Proceedings of 60th Annual Associated Schools* 5 (2024) 867–875.
- [4] S. G. Wheeler, H. Engelbrecht, S. Hoermann, Human factors research in immersive virtual reality firefighter training: A systematic review, *Frontiers in Virtual Reality* 2 (2021) 671664.

- [5] M. Porubcinova, H. Fidlerova, Systematic review of virtual reality tools for firefighting training based on the practical experience of a firefighter, *Information Technologies and Learning Tools* 108 (2025) 64.
- [6] L. O. Gomes, P. L. Oliveira, P. S. Mattos, M. Sacks, Enhancing firefighter training through virtual reality: Evaluating performance gains, user acceptance, and the impact of prior experience, *training* 16 (2026) 17.
- [7] J. Hofmann, S. Polikarpus, Application of Augmented Reality in firefighters training: From Safe to SafAR, in: *GI VR/AR Workshop, Gesellschaft für Informatik e.V.*, 2022. doi:10.18420/vrar2022_184.
- [8] C. Papakostas, C. Troussas, A. Krouska, C. Sgouropoulou, Measuring user experience, usability and interactivity of a personalized mobile augmented reality training system, *Sensors* 21 (2021) 3888.
- [9] M. Sužnjević, K. Mišura, S. Srebot, L. Skorin-Kapov, M. Matijašević, From Training to Incident Response: Resilient XR and Digital Twin Support in Firefighting, in: *Proceedings of the 1st Workshop on XR for Challenging Environments (XR4CE)*, co-located with ACM CHI 2026, *CEUR Workshop Proceedings*, 2026.
- [10] T. Franke, C. Attig, D. Wessel, A personal resource for technology interaction: development and validation of the affinity for technology interaction (ati) scale, *International Journal of Human-Computer Interaction* 35 (2019) 456–467.
- [11] R. S. Kennedy, N. E. Lane, K. S. Berbaum, M. G. Lilienthal, Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness, *The international journal of aviation psychology* 3 (1993) 203–220.
- [12] J. Brooke, et al., Sus-a quick and dirty usability scale, *Usability evaluation in industry* 189 (1996) 4–7.
- [13] J. C. Byers, A. Bittner, S. G. Hill, Traditional and raw task load index (tlx) correlations: Are paired comparisons necessary, *Advances in industrial ergonomics and safety* 1 (1989) 481–485.
- [14] F. D. Davis, Perceived usefulness, perceived ease of use, and user acceptance of information technology, *MIS quarterly* (1989) 319–340.
- [15] F. D. Davis, R. P. Bagozzi, P. R. Warshaw, User acceptance of computer technology: A comparison of two theoretical models, *Management science* 35 (1989) 982–1003.