

Mixed Reality-Supported Training of First Responder Skills in International Crisis Situations: Evaluation of the SmartSkills Pilot Study

Georg Aumayr¹, Andreas Peer^{2,3}, Benjamin Schuster¹, Martin Müller¹, Christian Schönauer², Thomas Seirlehner², Sabrina Hackl², Stefan Ladstätter⁴, Anna Weber⁴, Michael Schneeberger⁴, Martin Pszeida⁴, Wolfgang Weiss⁴, Florian Haid⁴, Jochen A. Mosbacher⁴, Joachim Brandtner⁵, Patrick Luley⁴, Markus Bergen⁴, Friedrich Urbanek², K. Wolfgang Kallus⁶, Markus Öttl³, Gudrun Walter³, and Lucas Paletta⁵

¹Johanniter Österreich Ausbildung und Forschung gemeinnützige GmbH, Vienna, Austria

²M2D MasterMind Development GmbH, Langenlebarn, Austria

³Austrian Armed Forces, Austria

⁴JOANNEUM RESEARCH Forschungsgesellschaft mbH, Graz, Austria

⁵Bytewood GmbH, Vienna, Austria

⁶Institut für Begleitforschung und Psychologisches Qualitätsmanagement, Graz, Austria

ABSTRACT

Military first responders and civilian experts in international peace missions must make rapid, safety-critical decisions while managing conflict escalation under acute stress. Live field exercises such as Native Challenge provide high realism but are costly, resource-intensive and difficult to repeat in a standardized way. The Austrian SmartSkills project addresses this gap by implementing Mixed Reality (MR) training that combines repeatable crisis scenarios with high-fidelity digital twins, instructor-facing decision support and human-factors monitoring. This paper presents the design and pilot evaluation of the Illegal Checkpoint scenario at the Johanniter Simulation Centre in Vienna, comparing MR-supported training with a conventional real simulated reference. The evaluation combines qualitative self-judgement, trainer and expert ratings of leadership and stress coping, standardized questionnaires on computer anxiety (Heinssen et al., 1987), technology acceptance (Venkatesh & Bala, 2008), immersion (Jennett et al., 2008), self-efficacy (Murphy et al., 1989; Compeau & Higgins, 1995) and learning preferences (Fleming & Mills, 1992; Kolb, 1984; Kolb & Kolb, 2013), system logs, and selected event-synchronized physiological examples. Results indicate that live training remains superior in embodied social intensity, but that MR training can approximate key stress reactions, provide high immersion when technically stable, and substantially reduce staffing and material requirements. The contribution is a structured evaluation framework for repeatable, evidence-informed crisis training without claiming that virtual simulation substitutes live exercises.

Keywords: Skill training, Mixed reality, Peace missions, Conflict scenarios, Biosignal sensors, Pilot study, Digital twins, Questionnaires.

INTRODUCTION

Military first responders, humanitarian teams and civilian experts in peace missions must act under uncertainty, social threat and acute time pressure. In illegal checkpoint situations, the field team must maintain de-escalating communication, comply with hostile instructions, protect team members and avoid actions that could trigger escalation. These skills are difficult to train repeatedly in live exercises because realistic role-player scenarios require staff, space, equipment, safety procedures and time-intensive debriefing.

The SmartSkills project transfers central components of crisis-operation training into a Mixed Reality framework. The intention is not to replace live exercises, but to create a repeatable intermediate training layer between classroom instruction and costly field exposure. MR can preserve spatial and social realism through digital twins and role-player interaction, while adding scenario control, event logging, questionnaires and selected human-factors data for structured after-action review.

For user organisations, immersive training environments offer a practical advantage: basic training and selected advanced drills can be performed in a secured environment with less preparation time, reduced staffing and lower material consumption. Real exercises often operate as one-shot experiences in which all logistics and role-player structures must be available at once. Such exercises can create strong “learning by shock” effects, but deeper cognition and stress coping require repetition. A standardized simulation enables repeated exposure to relevant scenario parts, targeted variation of stressors, and focused debriefing without the full cost of a real-life setup.

The research questions were: (1) Can the Illegal Checkpoint scenario be transferred into a digital-twin MR setting without losing relevant operational affordances? (2) Is the MR system sufficiently robust for instructor-led simulation-centre use? (3) How do trainees and trainers evaluate MR training compared with a conventional real simulated setup in terms of realism, workload, acceptance and learning value? (4) Can scenario events, questionnaires and selected physiological examples be synchronized to support future debriefing?

RELATED WORK

MR technologies are increasingly used for training in domains where field errors may have severe consequences and live repetition is expensive or unsafe. Recent reviews and empirical studies report benefits for emergency and medical first responders, including triage accuracy, perceived learning, and feasibility of immersive simulation (Cardós-Alonso et al., 2024; Mühling et al., 2023). MR training value depends not only on visual immersion but also on interaction, social presence, instructor control and user acceptance (Uhl et al., 2025; Zechner et al., 2024).

The digital-twin component addresses the requirement that training scenarios remain spatially credible and functionally equivalent to reference environments. Digital twins are virtual representations connected to physical assets or processes (Negri et al., 2017), and recent work on MR and digital twins demonstrates their value for safe, repeatable interaction with complex

environments (Piñal and Arguelles, 2024). For SmartSkills, the digital twin must represent operational affordances such as visibility, blocked routes, cover, actor approach paths and object interaction.

Evaluation of MR training must combine system, human-factors and learning perspectives. Situation awareness remains critical because trainees must perceive relevant objects and actors, comprehend their meaning and anticipate near-future developments (Endsley, 1995). Stress-probe concepts draw on laboratory findings that social-evaluative threat and uncontrollability are central stressor features (Dickerson and Kemeny, 2004). Pupillometry and cardiovascular measures can index arousal and regulation when interpreted with scenario context (Wang et al., 2018; Widmann et al., 2018).

NATIVE CHALLENGE AND ILLEGAL CHECKPOINT SCENARIO

The Native Challenge is a realistic peace-mission exercise framework in which participants experience simulated conflict scenarios as members of a field team. The illegal checkpoint scenario combines spatial constraint, social threat and time pressure. A team approaching a blocked road is confronted by armed role players who may demand documents, separate team members, seize communication devices, search bags or vehicles and increase hostility depending on trainee behaviour.

The training goal is safe de-escalation rather than tactical dominance. Trainees have to remain calm, communicate cooperatively, comply with instructions, monitor the team and avoid unnecessary negotiation or contradiction. The scenario is therefore well suited for MR-supported training: spatial structure can be standardized through a digital twin, while human role players preserve social unpredictability, proximity and emotional credibility.



Figure 1: Reference scenes for SmartSkills training: (a) illegal checkpoint confrontation from the Native Challenge field setting and (b) crisis-operation casualty-care context with MR-based realistic view from the simulation-centre trial. The scenes define the operational realism that the pilot aims to preserve while improving repeatability and measurement precision.

OVERALL SYSTEM ARCHITECTURE

The SmartSkills system consists of four connected layers (Figure 2). The first layer is the digital-twin and scenario-authoring layer, where spatial captures are transformed into optimized 3D assets and enriched with objects, hazards, actors and event markers. The second layer is the MR training layer, implemented with a Meta Quest 3-based SmartSkills application and physical space for trainees, trainers and role players. The third layer is the data layer, which synchronizes MR logs, scenario markers, questionnaire metadata and selected biosignal streams. The fourth layer is the trainer-support layer, which provides dashboards and replay material for debriefing.

The architecture deliberately supports different levels of instrumentation. The pilot can be evaluated even if physiological data are incomplete, because the main evidence comes from scenario execution, digital-twin review, questionnaires, trainer observation and system logs. Biosignals are integrated as optional explanatory signals around marked events, not as independent diagnostic labels.

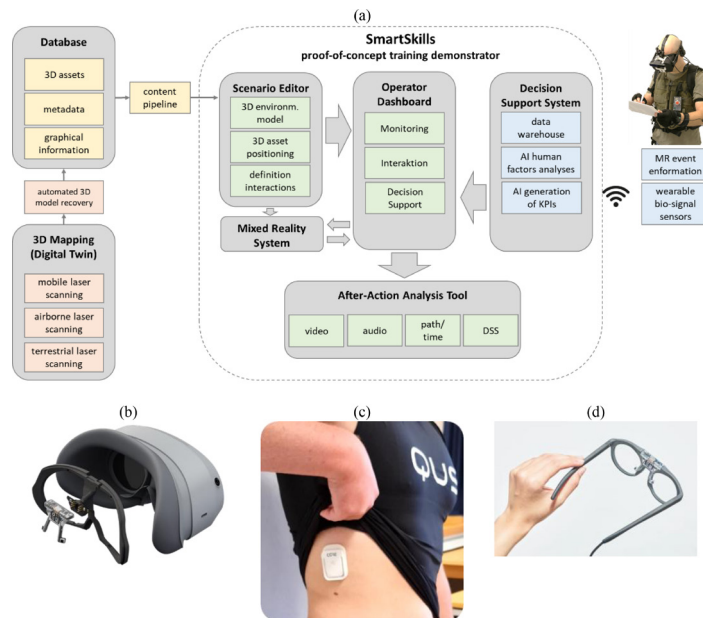


Figure 2: SmartSkills system architecture and selected components: (a) digital twin, human-factors analytics, MR simulator and 3D visualization pipeline; (b) MR headset with eye-tracking embedding; (c) smart textile cardiovascular unit; and (d) mobile eye-tracking glasses for field data collection.

3D ASSET AND CHARACTER DEVELOPMENT FOR THE SMARTSKILLS TRAINING ENVIRONMENT

The SmartSkills project required a comprehensive library of high-fidelity 3D assets and character models to populate the digitally reconstructed training environment. A dedicated production pipeline ensured that assets met the

visual and spatial fidelity requirements of a photorealistic MR training system while remaining optimized for real-time rendering on the target platform. Asset production used structured photographic reference acquisition of objects, equipment and personnel under controlled conditions. This material guided geometric modelling and physically plausible surface authoring so that virtual assets remained recognizable and behaviourally consistent with operational counterparts.

Where geometric complexity warranted it, photogrammetric scan data served as a baseline for production-ready meshes derived through manual retopology. This preserved spatial accuracy while satisfying real-time polygon budgets. For assets captured through photographic reference, modelling was performed from scratch in Blender with attention to silhouette accuracy and surface detail at the scale at which trainees perceive objects during scenario execution.

The asset catalogue covered military load-bearing gear, tactical backpacks and weapons as well as emergency medical props such as roller bandages, scoop stretchers and items representative of the Johanniter emergency response toolkit. Each prop was modelled, UV-unwrapped and textured to production standards, with albedo, normal and roughness maps authored in Adobe Photoshop and baked from high-resolution source geometry. Character development included military personnel and Johanniter emergency responders as fully rigged and animated male and female variants. Uniform markings, equipment attachment points and role identifiers were represented in surface materials. Skeletal rigs supported casualty handling, equipment deployment and coordinated team movement. The completed asset library was integrated into the digital-twin environment, while the environmental reconstruction track focused on architectural and infrastructural spatial accuracy. Together, both streams addressed perceptual fidelity, operational representativeness and real-time performance.

INTEGRATION OF PHYSICAL PROPS IN MIXED REALITY

SmartSkills uses two complementary MR approaches to integrate real objects into the digital-twin environment. Selected physical props can be equipped with Antilatency trackers, allowing their position and rotation to be represented continuously in the virtual scene. In addition, real objects can be placed in a defined passthrough region, implemented as a greenscreen-style window on the floor, where the Quest camera image remains visible inside the otherwise virtual environment. Together, these approaches combine spatially tracked interactive props with untracked real-world objects that remain visible and usable in a controlled physical area.

The tracking setup combines inside-out tracking of the Meta Quest 3 with an Antilatency optical-inertial layer. Quest tracking provides headset pose and hand interaction, while Antilatency adds persistent 6DoF references for the trainee's lower body and selected props. In Unity, Antilatency support is encapsulated in a prefab containing the device network, configured tracking environment, head and leg tracker transforms, and numbered tracker transforms for physical objects. The project uses Antilatency SDK

4.1.0 together with Meta XR, Unity XR Hands, Oculus XR and Unity Netcode. At runtime, the handler links tracked head and leg objects to trainee synchronization objects and compensates horizontal drift between the Antilatency head reference and Quest camera pose. If a threshold is exceeded, the trainee rig is translated and yaw-corrected so both coordinate frames remain aligned (Figure 5).

Before scenario execution, the trainee performs a short calibration pose. Standing upright in a T-pose synchronizes avatar scale, head position and leg offsets. After this synchronization, the trainee can switch from controllers to hand tracking. Avatar animation uses a VRIK-based synchronization pipeline: Quest tracking drives head and hands, while Antilatency leg trackers act as additional lower-body targets when enabled. If Antilatency is disabled, leg IK weights are removed and the avatar returns to standard Quest-based tracking. Scenario objects can be spawned and assigned to numbered tracker tags, and the assignment is saved with the scene to support repeatability and after-action review.

MIXED REALITY VISUALIZATION AND TRAINER SUPERVISION

The SmartSkills MR visualization layer supports greenscreen-style compositing to integrate selected real-world areas into the virtual training scene. In the pilot setup, a passthrough “window” on the floor covered by a green cloth allows real objects on the cloth to remain visible to the trainee while the surrounding environment is rendered as the digital-twin scenario. Meta Quest passthrough and chroma-key-style compositing settings assign the live camera image to predefined scene geometry such as floor windows, doors, cubes or other surfaces, while the rest of the scene remains virtual.

A second visualization mode supports trainer supervision. A trainer or adversary role player can observe the scenario through a Quest headset with adjustable transparency of the virtual world. By changing transparency, the trainer sees both virtual scenario content and real trainees moving in the physical training space. This supports safety monitoring of distances, posture, obstacles and role-player interaction, and it improves instructional control because trainers can judge how virtual events, real actions and social interaction unfold together.

3D RECONSTRUCTED DIGITAL TWIN ENVIRONMENT

The digital twin of the Alpine training environment was generated using a multi-sensor reality-capture workflow combining aerial photogrammetry and terrestrial laser scanning (TLS). A camera-equipped unmanned aerial vehicle (UAV) was used to acquire high-overlap imagery of the training environment, providing complete coverage of roofs, open terrain, vegetation, and inaccessible areas. In parallel, a terrestrial laser scanning system captured high-density three-dimensional measurements of buildings, infrastructure, and other scene elements from ground level, ensuring accurate reconstruction of vertical structures and occluded regions.

The digital twin was generated from a multi-sensor reality-capture workflow combining airborne and terrestrial LiDAR mapping with high-resolution image acquisition. Data collection was performed using a RIEGL RiCOPTER VUX-SYS airborne laser scanning platform and a RIEGL VZ-400i terrestrial laser scanner. Both systems integrate precisely calibrated cameras and tightly coupled GNSS/IMU navigation units, enabling direct georeferencing of the acquired data and accurate co-registration of point clouds and imagery without the need for extensive photogrammetric alignment procedures.

The airborne and terrestrial LiDAR datasets were processed to generate a geometrically accurate and complete three-dimensional representation of the training environment. The complementary acquisition geometries provided detailed coverage of both horizontal and vertical surfaces while minimizing occlusions commonly encountered in single-platform surveys. Because both mapping systems produce accurately georeferenced point clouds, the individual datasets could be merged directly into a common coordinate framework, resulting in a dense and spatially consistent representation of the scene.



Figure 3: LiDAR mapping platforms which were used for collecting measurement data: Ricopter VUX-SYS (left) and Riegl VZ-400i (right).



Figure 4: 3D Reconstruction of training area. Detailed view (left) and total area (right).

The integrated camera systems were used to colorize the LiDAR point clouds by projecting image information onto the measured points using the calibrated camera poses and orientation data. In addition, the high-resolution imagery was employed to enhance the visual quality of the final digital twin. While the geometric reconstruction relied primarily on the LiDAR measurements, photogrammetric image-processing techniques were used to derive high-resolution texture information that exceeded the spatial sampling density achievable from the laser measurements alone. This approach combined the geometric accuracy and robustness of LiDAR with the visual richness of image-based reconstruction.

Following data fusion, the point clouds were filtered and cleaned to remove noise and transient objects before being converted into textured surface meshes. Semi-automated meshing algorithms were used to reconstruct continuous surfaces from the point cloud data, followed by topology optimization and mesh simplification to achieve real-time rendering performance. High-resolution texture maps were generated from the image dataset and projected onto the reconstructed geometry, producing a photorealistic representation of the environment. Multiple levels of detail (LOD) were subsequently created to support efficient visualization on Mixed Reality hardware.

The resulting digital twin provides a geometrically accurate and visually realistic representation of the training area while preserving operationally relevant spatial affordances such as visibility corridors, blocked routes, cover positions, actor approach paths, and interaction-relevant objects. This combination of LiDAR-based geometric reconstruction and image-based texture enhancement enabled the creation of simulation-ready 3D assets suitable for immersive Mixed Reality training scenarios.

PILOT STUDY EVALUATION

The exploratory pilot was conducted at the Johanniter Simulation Centre in Vienna in 2025 as a proof-of-concept evaluation of the SmartSkills training ecosystem (Figure 5). Twelve participants took part (10 males, 2 females; 34.33 +/- 12.78 years of age; Table 1). Most participants reported no prior VR experience. The study comprised six trials; each trial included one trainer and two trainees and combined briefing, baseline or familiarization, scenario execution, debriefing and questionnaires (Table 2). The MR condition used the digital-twin environment and the SmartSkills application while retaining human trainers and role players. The conventional condition served as a live-reference simulation with the same objective and comparable role-player pressure.



Figure 5: Implementation and debriefing context of the Vienna pilot: (a) wearable setup, (b) trainee in the MR simulation environment validating external calibration targets, (c) synchronized dashboard context, and (d) mixed-reality training space with real actors and displays.

In the real-life test condition, the experience initially contained typical exercise artefacts and was sometimes perceived as artificial or even amusing. Within seconds, however, the immersion increased and trainers took over control, running the scenario as if it were real. The physical setup addressed all senses and added embodied impact. For the conventional scenario, three trainers and one scripted insider role player had to be staffed for each participant group in order to provide the full interaction logic. In the XR scenario, one trainer was sufficient.

Table 1: Socio-biographic information about the pilot participants.

Variable	Mean	SD	N
Age	34.33	12.78	9
Highest education level	2.00	1.00	9
MR experience	1.88	0.35	8
Technology use	22.11	4.40	9
Anxiety	26.00	2.18	9

Qualitative evaluation combined participant self-judgement, trainer feedback and expert ratings for leadership and stress coping. Quantitative data were collected using a standardized questionnaire battery comprising elements from the Computer Anxiety Rating Scale (CARS), Technology Acceptance Model III (TAM III), Immersive Experience Questionnaire (IEQ), Computer Self-Efficacy Scale, VARK Questionnaire and Kolb Learning Style Inventory. Additional system logs documented calibration, tracking behaviour, technical interruptions and event timing (Table 2).

Table 2: Overview on instruments used for the pilot assessment.

Assessment Construct	Instrument Basis	Elements Used in SmartSkills	Score Interval Used/ Interpreted
Computer anxiety	CARS	Computer-related anxiety and nervousness before/around technology use	CARS full scale typically 19-95; in the pilot analysed as summed anxiety score
Technology acceptance	TAM III	Perceived usefulness, usability/ ease of use, enjoyment, result demonstrability/results	Subscales interpreted as summed Likert scores; key acceptance subscales in the pilot were typically 6-30
Self-efficacy	Computer Self-Efficacy Scale / TAM III anchor	Confidence in being able to use the system and perform with it	Short subscale interpreted approximately 3-15
Immersion	IEQ	Absorption, involvement and dissociation from the real surroundings	Adapted short immersion score, approximately 8-40

(Continued)

Table 2: Continued.

Assessment Construct	Instrument Basis	Elements Used in SmartSkills	Score Interval Used/ Interpreted
Learning preferences	VARK; Kolb LSI	Visual, Aural, Read/Write, Kinesthetic preferences; experiential learning modes CE, RO, AC, AE	Profile/count scores, not a single total score; used descriptively
Presence and situation awareness	Presence questionnaire; SART-style SA	Sense of being there; understanding, attentional demand and attentional supply	Presence reported as summed score; SA reported as summed SART-style score

QUESTIONNAIRE AND CORRELATIONAL RESULTS

Questionnaire availability varied across constructs. Most correlation analyses used $N = 9$ valid participants, with $N = 8$ for VR/MR experience and $N = 6$ for additional future-use questions. Results are therefore exploratory and should be interpreted as effect-size indications, not confirmatory hypothesis tests (Table 1).

The strongest sociodemographic association was between education and anxiety ($r = -.688$, $p = .040$), suggesting that the coded education variable was inversely related to computer anxiety. In the simulated exercise (SE) block, technology use correlated with perceived usefulness ($r = .720$, $p = .029$). SE perceived usefulness was strongly related to usability ($r = .671$, $p = .048$), enjoyment ($r = .960$, $p < .001$) and perceived results ($r = .786$, $p = .012$). Usability was also associated with enjoyment ($r = .804$, $p = .009$). This pattern indicates a coherent acceptance profile: participants who considered the system useful also tended to experience it as usable, enjoyable and effective.

In the MR block (Table 3), the clearest significant relation was between enjoyment and perceived results ($r = .695$, $p = .038$). Several associations were close to conventional significance and are useful for hypothesis generation: technology use and MR situation awareness ($r = .636$, $p = .066$), perceived usefulness and presence ($r = .659$, $p = .075$), usability and presence ($r = .615$, $p = .078$), and MR results and situation awareness ($r = .633$, $p = .067$).

Additional future-use ratings were available for $N = 6$. For MR, visual-display danger showed a trend-level negative association with perceived usefulness for future training ($r = -.745$, $p = .089$).

Table 3: MR experience of the participants captured with a mixed questionnaire.

MR Variable	Mean	SD	N
Perceived usefulness	24.63	3.38	8
Usability	24.67	3.43	9
Enjoyment	20.11	1.54	9
Results	26.00	3.04	9

(Continued)

Table 3: Continued.

MR Variable	Mean	SD	N
Self-efficacy	9.33	2.65	9
Immersion	33.89	3.69	9
Presence	84.89	12.97	9
Situation awareness	27.22	5.97	9

QUALITATIVE RESULTS AND TRAINING VALUE

Qualitative feedback from trainers and trainees favoured the real physical training reference, especially for embodied intensity, social pressure and the credibility of physical interaction. However, the MR training was not evaluated as far behind the real condition. Participants valued the general immersion highly until technical disruptions occurred. The main issues in the MR condition were moving artefacts and strange behaviour of non-player characters, which occasionally interrupted attention and reduced focus. These problems were interpreted as implementation issues rather than conceptual objections to MR-based training.

Stress-related reactions were broadly comparable in their direction: trainees responded physiologically and behaviourally to stressors in both the real and MR settings. Interestingly, the physical impact of the real-life training was evaluated as less intense in subjective reception than the MR scenario in some comments, possibly because the virtual setup concentrated attention on specific stress cues. Overall, the MR setup appears promising for delivering adequate training opportunities with significant cost reduction. It should not be presented as a substitute for real-life exercise and drill, but as a repeatable preparatory and complementary training layer.

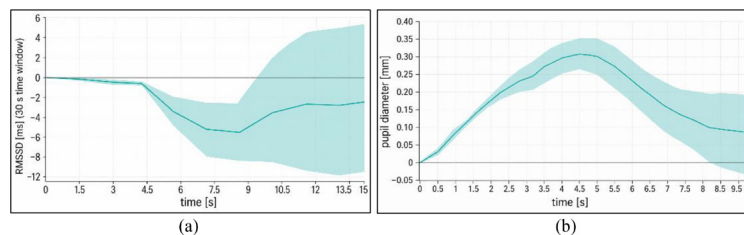


Figure 6: Event-related physiological responses to a standardized MR stress probe: predefined digital stimulus (aggressive shouting) was presented as a short auditory-social stressor and used as a time-locked event marker. (a) RMSSD response calculated across 30-s sliding windows, illustrating short-term cardiovascular regulation during and after probe exposure; (b) pupil-diameter response in the corresponding event-aligned window, illustrating rapid pupillometric arousal dynamics elicited by the stress probe (from Schneeberger et al., in press).

BIOSIGNAL-SENSOR-BASED STRESS MEASUREMENTS

Biosignal sensing was used as a compact supporting layer. MR-based data from the Meta Quest 3 SmartSkills app, eye-tracking information and selected cardiovascular signals were transmitted to a central data warehouse and synchronized with scenario markers. Event windows around stress probes can be used to inspect heart-rate changes, HRV/RMSSD variation, pupil-dilation responses and gaze behaviour. The aim is not diagnostic classification, but debriefing support: trainers can identify episodes worth replaying, such as delayed compliance, missed peripheral cues or high workload around specific probes.

Furthermore, HRV/RMSSD and pupillometry can be used as event-synchronized stress indicators during short, standardized stress probes in MR training (Schneeberger et al., in press). RMSSD provides information about short-term parasympathetic regulation and recovery, while pupil diameter reflects rapid autonomic arousal and cognitive-emotional activation (Figure 6). When aligned to defined probe onset times, these signals can help trainers identify moments of increased strain, compare responses across repetitions, and support focused debriefing without replacing expert judgement.

DISCUSSION

The SmartSkills pilot reframes MR crisis training as a socio-technical evaluation problem. The value of the system does not depend only on headset immersion or sensor availability, but on the alignment between digital-twin realism, human role-player credibility, trainer control, event logging and debriefing. A technically impressive MR scene would be insufficient if trainees did not experience relevant pressure, if trainers could not control the interaction or if debriefing did not connect observed behaviour to learning goals.

The comparison with a conventional real simulated setup is essential. Live simulation remains the reference for social intensity and embodied interaction. MR can add repetition, safe variation, standardized stress probes and structured data capture. The expected advantage is not that MR is more real than live training, but that it can prepare trainees for live exposure by allowing controlled repetition and targeted feedback on de-escalation, compliance and team coordination. The cost-benefit argument is especially relevant for user organisations because virtual scenarios reduce staffing and material requirements and support repeated practice of drill components that are hard to isolate in full-scale exercises.

Biosignal sensing should be interpreted conservatively. HR, HRV, pupillometry and gaze can support debriefing when aligned to scenario events, but they are influenced by movement, breathing, posture, luminance, headset fit and individual baselines. In SmartSkills, the near-term value of biosignals is therefore episode selection rather than automatic labels of trainee competence.

LIMITATIONS AND FUTURE WORK

The pilot is exploratory, with a small sample and heterogeneous data availability. It is not designed to establish diagnostic thresholds, prove equivalence between MR and live simulation or quantify long-term transfer. Questionnaire results may be influenced by novelty effects, prior VR experience and differences in role-player behaviour. System results may also depend on local room layout, headset calibration, network conditions and trainer familiarity with the dashboard.

Future work should extend the sample, standardize the probe library, compare repeated MR sessions with live-reference training and assess retention of de-escalation skills.

CONCLUSIONS

SmartSkills advances MR first-responder training from immersive visualization toward structured, evidence-informed learning analytics. The Vienna pilot places the complete training ecosystem at the centre: digital twins, MR interaction, human role players, physical props, stress probes, questionnaires, system logs, debriefing and selected biosignal examples.

Results indicate that real exercises remain indispensable for embodied drill and full social intensity, but that MR training can provide a repeatable, cost-efficient and instructor-controllable preparation layer. The framework supports repeatable, ethically controllable and trainer-centred crisis training while preserving the social realism needed for checkpoint de-escalation.

ACKNOWLEDGMENT

This work was partially funded by the projects SmartSkills (grant No. FO999905346) and SmartResponse (grant No. FO999926240) of the Austrian FFG KIRAS programme, co-financed by the Austrian Federal Ministry for Innovation, Mobility and Infrastructure (BMIMI), as well as by the project iMEDCAP, co-financed by the European Union under Grant Agreement No: 101121421. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Commission. Neither the European Union nor the granting authority can be held responsible for them.

REFERENCES

- Cardós-Alonso, M. C., Otero-Varela, L., Redondo, M., and González-García, A. (2024). Extended reality training for mass casualty incidents: a systematic review on effectiveness and experience of medical first responders. *International Journal of Emergency Medicine*, 17, 112. <https://doi.org/10.1186/s12245-024-00685-3>.
- Compeau, D. R., and Higgins, C. A. (1995). Computer self-efficacy: Development of a measure and initial test. *MIS Quarterly*, 19(2), 189–211. <https://doi.org/10.2307/249688>.

- Dickerson, S. S., and Kemeny, M. E. (2004). Acute stressors and cortisol responses: a theoretical integration and synthesis of laboratory research. *Psychological Bulletin*, 130(3), 355-391. <https://doi.org/10.1037/0033-2909.130.3.355>.
- Endsley, M. R. (1995). Toward a theory of situation awareness in dynamic systems. *Human Factors*, 37(1), 32-64. <https://doi.org/10.1518/001872095779049543>.
- Fleming, N. D., and Mills, C. (1992). Not another inventory, rather a catalyst for reflection. *To Improve the Academy*, 11(1), 137-155. <https://doi.org/10.1002/j.2334-4822.1992.tb00213.x>.
- Heinssen, R. K., Glass, C. R., and Knight, L. A. (1987). Assessing computer anxiety: Development and validation of the Computer Anxiety Rating Scale. *Computers in Human Behavior*, 3(1), 49-59. [https://doi.org/10.1016/0747-5632\(87\)90010-0](https://doi.org/10.1016/0747-5632(87)90010-0).
- Jennett, C., Cox, A. L., Cairns, P., Dhoparee, S., Epps, A., Tijs, T., and Walton, A. (2008). Measuring and defining the experience of immersion in games. *Intl Journal of Human-Computer Studies*, 66(9), 641-661. <https://doi.org/10.1016/j.ijhcs.2008.04.004>.
- Kolb, D. A. (1984). *Experiential learning: Experience as the source of learning and development*. Prentice Hall.
- Kolb, A. Y., & Kolb, D. A. (2013). The Kolb Learning Style Inventory 4.0: Guide to theory, psychometrics, research and applications. Experience Based Learning Systems.
- Murphy, C. A., Coover, D., & Owen, S. V. (1989). Development and validation of the Computer Self-Efficacy Scale. *Educational and Psychological Measurement*, 49(4), 893-899. <https://doi.org/10.1177/001316448904900412>
- Mühling, T., Späth, I., Backhaus, J., Milke, N., Oberdörfer, S., Meining, A., Latoschik, M. E., and König, S. (2023). Virtual reality in medical emergencies training: benefits, perceived stress, and learning success. *Multimedia Systems*, 29, 2239-2252. <https://doi.org/10.1007/s00530-023-01102-0>.
- Negri, E., Fumagalli, L., and Macchi, M. (2017). A review of the roles of Digital Twin in CPS-based production systems. *Procedia Manufacturing*, 11, 939-948. <https://doi.org/10.1016/j.promfg.2017.07.198>.
- Paletta, L., Peer, A., Öttl, M., Aumayr, G., Kallus, K., Brandtner, J., Walter, G., Luley, P., Almer, A., Weismeier-Sammer, D., Schönauer, C., Schuster, B., Söllner, M., Müller, M., Schneeberger, M., Pszeida, M., Weiss, W., Weber, A., Ladstätter, S., Haid, F., Mosbacher, J., Russegger, S., and Bergen, M. (2025). A framework for Mixed Reality-supported training of conflict resolution and first responder skills in international crisis situations: SmartSkills. In L. Paletta (Ed.), *Cognitive Computing and Internet of Things*. AHFE Open Access, 165. <https://doi.org/10.54941/ahfe1006097>.
- Piñal, O., and Arguelles, A. (2024). Mixed reality and digital twins for astronaut training. *Acta Astronautica*, 218, 163-178. <https://doi.org/10.1016/j.actaastro.2024.01.034>.
- Schneeberger, M., Pszeida, M., Weiss, W., Peer, A., Aumayr, G., Mosbacher, J. A., Haid, F., Luley, P., Weber, A., Ladstätter, S., Brandtner, J., Schönauer, C., Schuster, B., Söllner, M., Öttl, M., Walter, G., Kallus, K. W., & Paletta, L. (in press). Wearable biosignal and pupillometry analytics for stress-probe evaluation in mixed reality illegal checkpoint training. In L. Paletta (Ed.), *Cognitive computing in sociotechnical systems*. AHFE Open Access.
- Uhl, J. C., Zechner, O., Baetzner, A., Birrenbach, T., Egger-Lampl, S., Schrom-Feiertag, H., and Tscheligi, M. (2025). Mixed reality training for medical first responders: system evaluation and recommendations. *Virtual Reality*, 29, 69. <https://doi.org/10.1007/s10055-025-01144-x>.

- Venkatesh, V., and Bala, H. (2008). Technology Acceptance Model 3 and a research agenda on interventions. *Decision Sciences*, 39(2), 273-315. <https://doi.org/10.1111/j.1540-5915.2008.00192.x>.
- Wang, C.-A., Baird, T., Huang, J., Coutinho, J. D., Brien, D. C., and Munoz, D. P. (2018). Arousal effects on pupil size, heart rate, and skin conductance in an emotional face task. *Frontiers in Neurology*, 9, 1029. <https://doi.org/10.3389/fneur.2018.01029>.
- Widmann, A., Schröger, E., and Wetzel, N. (2018). Emotion lies in the eye of the listener: emotional arousal to novel sounds is reflected in the sympathetic contribution to the pupil dilation response and the P3. *Biological Psychology*, 133, 10-17. <https://doi.org/10.1016/j.biopsycho.2018.01.010>.
- Zechner, O., Schrom-Feiertag, H., Wespi, R., Pretolesi, D., Nguyen, Q., and Tscheligi, M. (2024). Enhancing mixed reality simulation training technology with real-time performance visualization: mixed methods study with medical first responders. *JMIR XR and Spatial Computing*, 1, e57655. <https://doi.org/10.2196/57655>.