Cognitive Computing in Sociotechnical Systems, Vol. 205, 2026, 234–248
https://doi.org/10.54941/ahfe1007377

Wearable Biosignal and Pupillometry Analytics for Stress-Probe Evaluation in Mixed Reality Illegal Checkpoint Training

Michael Schneeberger[1], Martin Pszeida[1], Wolfgang Weiss[1], Andreas Peer[2,3], Georg Aumayr[4], Jochen A. Mosbacher[1], Florian Haid[1], Patrick Luley[1], Anna Weber[1], Stefan Ladstätter[1], Joachim Brandtner[5], Christian Schönauer[2], Benjamin Schuster[4], Martin Söllner[4], Markus Öttl[3], Gudrun Walter[3], K. Wolfgang Kallus[6], and Lucas Paletta[1]

[1]Joanneum Research Forschungsgesellschaft mbH, Graz, Austria
[2]M2D MasterMind Development GmbH, Langenlebarn, Austria
[3]Austrian Armed Forces, Austria
[4]Johanniter Österreich Ausbildung und Forschung gemeinnützige GmbH, Vienna, Austria
[5]Bytewood GmbH, Vienna, Austria
[6]Institut für Begleitforschung und Psychologisches Qualitätsmanagement, Graz, Austria

ABSTRACT

High-stakes peace mission and first-responder training must expose trainees to threat, ambiguity and time pressure without compromising safety or repeatability. This paper presents a cutting-edge Mixed Reality (MR) training and learning-analytics concept for the novel “SmartSkills” training procedure and reports an exploratory proof-of-concept pilot study conducted at the Johanniter Simulation Centre in Vienna in 2025. The system combines digital twins of operational training sites, live and virtual actors, a scenario editor, eye tracking, cardiovascular and wearable biosignal sensing, and an instructor-facing decision support layer. The methodological core is a stress-probe approach: short, time-locked acoustic and interaction events, especially aggressive shouting and other emotionally aversive auditory cues, are embedded in unavoidable key scenes of an illegal checkpoint scenario. Physiological responses are analysed in standardized windows using heart rate, short-term HRV/RMSSD and pupillometry, with gaze-based indices of situation awareness. The pilot study included 12 participants in six trials, each involving a trainer and two trainees. Results indicate that synchronized wearable and eye-tracking data can identify probe-locked autonomic changes and longer episodes of sustained stress during MR training. Although the study is exploratory and not powered for inference, the approach provides a scalable pathway toward objective after-action review, adaptive scenario difficulty and individualized stress-resilience training.

Keywords: Mixed reality training, Stress probes, Pupillometry, Heart rate variability, Auditory stressors, Situation awareness, Digital twins

INTRODUCTION

Military first responders, humanitarian teams and civilian experts in peace missions must act under uncertainty, social threat, and acute time pressure.

Received June 14, 2026; Revised July 3, 2026; Accepted July 11, 2026; Available online July 20, 2026

Live exercises such as the Native Challenge in Tyrol provide rare high-fidelity exposure to illegal checkpoints, negotiation breakdowns, mine-awareness events and casualty care. However, live simulation is expensive, difficult to personalize, and limited in how often identical stressors can be repeated across trainees. The SmartSkills project (Paletta et al., 2025) addresses this gap by transferring core scenario elements into an MR framework that preserves operational realism while enabling repeatable and ethically controllable stress exposure.

This paper advances the original SmartSkills concept toward a measurement-driven training paradigm. Rather than treating physiological activation as a post-hoc curiosity, the system embeds predefined stress probes into the scenario logic. A stress probe is a short, marked event that is salient for the trainee, unavoidable in the task flow and precisely synchronized with biosignal data. In the illegal checkpoint scenario, the most relevant probes are auditory and social: abrupt hostile commands, aggressive shouting, threat gestures and sudden changes in role-player behaviour. These probes are used to compare physiological reactivity across trials, individuals and scenario versions.

In SmartSkills, biosignal-based stress assessment is therefore not merely a manipulation check, but a marker of training readiness: because socially evaluative, uncontrollable and time-critical stressors reliably evoke psychobiological stress responses (Kirschbaum et al., 1993; Dickerson & Kemeny, 2004; Kallus & Gaisbachgrabner, 2018; Schneeberger et al., 2022), repeated MR checkpoint probes can indicate whether trainees still show excessive autonomic activation, delayed recovery and unstable situation awareness, or whether stress exposure has already consolidated functional routines under pressure (Saunders et al., 1996; Driskell et al., 2001). This is particularly relevant for the Vienna pilot, where short aggressive acoustic events are used as operationally embedded probes, since emotionally arousing sounds modulate sympathetic pupil-dilation components (Widmann et al., 2018), while pupil size, heart rate and electrodermal activity jointly index autonomic arousal (Wang et al., 2018). The SmartSkills approach extends previous work on immersive first responder training, where psychophysiological measurements were used to quantify cognitive-emotional stress and to relate stress responses to situation-awareness risk (Schneeberger et al., 2022; Paletta et al., 2022).

The contribution is threefold. First, we provide an updated state-of-the-art framework for MR-based high-stakes training with biosignal and pupillometry analytics. Second, we formalize a stress-probe methodology for short auditory and socio-emotional stressors. Third, we report the experimental results regarding responses measured with biosignal sensors in the exploratory Vienna 2025 pilot study, which tested synchronization of wearable cardiovascular data, mobile eye tracking and MR events in a realistic illegal checkpoint training sequence.

RELATED WORK

Simulation-based training is increasingly used where errors in the field may have severe consequences. Disaster and emergency training studies have shown that virtual environments can support knowledge retention and scenario repetition when live exercises are impractical (Farra et al., 2013). Recent military work further indicates that virtual environments can reveal operationally relevant differences in visual search and situation awareness between trained soldiers and civilians (Enders et al., 2024). For SmartSkills, these findings motivate a hybrid MR approach: real bodies, real equipment and role players are retained, while scenario variability, digital twins and biosignal instrumentation provide repeatability and measurement precision.

Laboratory stress research has shown that robust psychophysiological stress responses are particularly likely when tasks combine social-evaluative threat and low controllability, two characteristics that are directly mirrored in illegal checkpoint situations where trainees are observed, challenged, separated, and forced to comply with hostile instructions (Dickerson & Kemeny, 2004). The Trier Social Stress Test further established the value of standardized, repeatable stress-induction procedures for eliciting measurable autonomic and endocrine responses under controlled conditions (Kirschbaum et al., 1993).

Situation awareness remains a central human-factors construct for checkpoint and crisis-response training. Endsley (1995) defines it as perception of relevant elements, comprehension of their meaning and projection of near-future states. In operational MR, situation awareness is not only a subjective rating but can be approximated by behavioural markers: gaze dwell time on threat-relevant objects, attention to peripheral actors, timely recognition of wounded persons or weapons, and compliance with de-escalation rules. This is consistent with distributed and team-oriented approaches to situation awareness that emphasize coupling between persons, artefacts and evolving environmental cues (Salmon et al., 2009).

Digital twins extend training realism by reconstructing sites, objects and environmental affordances as functional virtual counterparts. Reviews of digital-twin systems emphasize lifecycle representation, data fusion and interoperability between physical and virtual assets (Negri et al., 2017). For human-factors training, the relevant question is not only geometric realism but functional realism: Does a corner hide a threat? Does vehicle placement block retreat? Does the soundscape create urgency? The SmartSkills architecture therefore combines high-resolution spatial reconstruction with scenario semantics, event markers and AI-supported after-action analytics.

Biosignal-based stress analytics has matured from laboratory paradigms to wearable field sensing. HRV is a widely used marker of autonomic regulation, with RMSSD reflecting short-term parasympathetic activity when interpreted with appropriate windowing and artefact control (Task Force, 1996; Laborde et al., 2017; Shaffer and Ginsberg, 2017). Meta-analytic evidence shows that acute mental stress typically increases heart rate and reduces HRV, although effect sizes vary with task type, breathing, posture and measurement duration (Castaldo et al., 2015; Kim et al., 2018). EDA is particularly sensitive to sympathetic arousal and has been shown to help distinguish stress from cognitive load (Setz et al., 2010). Multimodal reviews conclude that robust stress detection benefits from combining cardiovascular, electrodermal, respiratory, movement and ocular features rather than relying on a single sensor channel (Giannakakis et al., 2022). Pupillometry is particularly relevant because pupil diameter is sensitive not only to visual processing, but also to autonomic arousal and cognitive-emotional engagement. Wang et al. (2018), for example, showed that pupil size can be interpreted together with heart rate and skin conductance as part of a multimodal arousal profile.

Pupillometry is especially attractive for MR and mobile field training because it captures arousal, cognitive effort and orienting responses at high temporal resolution. The pupil is influenced by luminance, gaze position and accommodation, but controlled or modelled conditions allow task-evoked dilation to be interpreted as an index of sympathetic and locus-coeruleus-noradrenergic activation. Emotional arousal increases pupil diameter during affective viewing (Bradley et al., 2008), while auditory studies demonstrate that pupil dilation is sensitive to listening effort, surprise and emotional arousal (Partala and Surakka, 2003; Zekveld et al., 2018; Montes-Lourido et al., 2021). Recent work further highlights that luminance, and arousal type must be modelled explicitly, which is critical in MR where displays and physical rooms jointly determine the pupil baseline (Pan et al., 2024).

Auditory stressors are particularly relevant for checkpoint training because they are fast, omnidirectional and difficult to ignore. Affective sound research shows that emotionally charged acoustic stimuli elicit systematic autonomic responses (Bradley and Lang, 2000; Gomez and Danuser, 2004; Chen et al., 2014). Human screams occupy a privileged acoustic niche characterized by roughness and strong salience for threat processing (Arnal et al., 2015). The Threat-of-Scream paradigm demonstrates that ethically acceptable aversive screams can be used to induce sustained subjective and physiological anxiety without electric shocks (Beaurenaut et al., 2020). In MR checkpoint scenarios, aggressive shouting, sudden commands and hostile vocal prosody provide ecologically valid short stressors while retaining experimental control.

Auditory stressors are especially suitable for short stress-probe paradigms because they can be presented abruptly, time-locked precisely, and embedded naturally in operational scenarios. Widmann et al. (2018) demonstrated that emotionally arousing novel sounds elicit pupil dilation responses with a sympathetic contribution, supporting the use of short aversive acoustic stimuli as measurable probes of autonomic orienting and arousal.

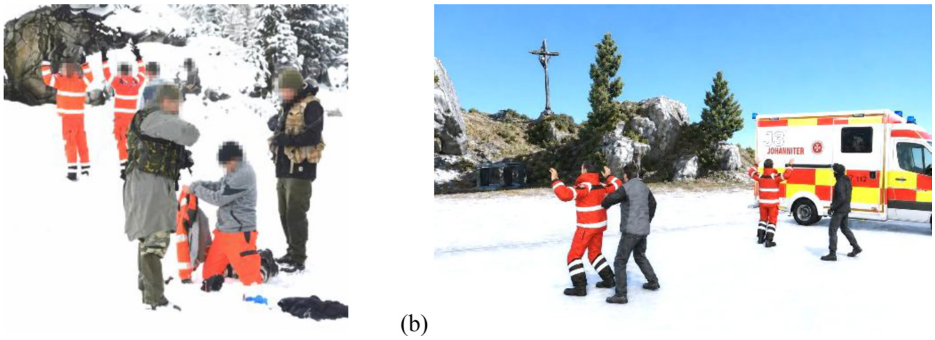
A key open challenge is therefore methodological: high-stakes training needs ecologically valid stressors, but learning analytics requires time-locked, repeatable events. The SmartSkills stress-probe concept addresses this by embedding short acoustic-social probes into a coherent scenario, synchronizing them with biosignals and using the same probes across real, MR and variant scenarios. This allows not only post-hoc interpretation but eventually adaptive training, where scenario intensity is modulated by the trainee's physiological response and behavioural performance.

NATIVE CHALLENGE AND ILLEGAL CHECKPOINT SCENARIO

The Native Challenge is a realistic peace-mission exercise framework in which students and professionals experience simulated conflict scenarios as members of a field team. One of the most demanding scenarios is the illegal checkpoint. The field team approaches a blocked road on foot or by vehicle, is confronted by armed role players and must avoid escalation. Training objectives include rapid threat appraisal, calm communication, de-escalation, compliance with hostile instructions, protection of team members and preservation of critical documents and communication equipment.

The scenario is well suited for MR because the functional structure is clear while the social dynamics can vary. The checkpoint is placed at a location where evasion is not possible; actors can separate trainees, demand identification, seize communication devices, search bags or vehicles and increase or decrease hostility in response to trainee behaviour.

The resulting training problem is not merely procedural. It requires emotion regulation, attention to peripheral cues and collective action under stress. Figure 1 anchors the MR concept in the live exercise that motivated the system.



(a) (b)

Figure 1: Fixed reference scene (illegal checkpoint) from the native challenge: (a) confrontation in the field (Tyrol, Austria) and (b) crisis-operation casualty care with MR-based realistic view from trial at the simulation centre. These scenes define the operational realism that the MR training system seeks to preserve while improving repeatability and measurement precision.

STRESS-PROBES FOR SHORT AUDITORY & SOCIAL STRESSORS

The stress-probe method defines short, marked scenario events that can be compared across trials. Each probe has a documented onset, duration and intensity class, is embedded at an unavoidable checkpoint of the task and is linked to an expected behavioural objective. In the Vienna pilot, short stress probes ("SP") lasted approximately 2-3 seconds, were digitally predefined and between one and three times repeatedly applied by purposeful control of external assistants. Longer stressor episodes ("ST") consisted of aggressive shouting or offensive comments lasting approximately 5-22 seconds, issued by the role-playing training staff as semi-intuitive part of the interaction plot.

The probes are not isolated laboratory tones; they are scenario-congruent events such as sudden hostile commands, intensified vocal prosody, weapon-related gestures or forced changes of posture.

The analysis principle is event related. Physiological data are synchronized to the probe onset, cleaned for artefacts and compared to a pre-probe baseline. Candidate features include heart-rate acceleration, RMSSD change, pupil dilation amplitude, time-to-peak, recovery slope, gaze reallocation toward threat sources and fixation coverage of task-relevant objects. Longer shouting episodes are evaluated as sustained stress phases using area-under-the-curve and recovery measures. Figure 2 summarizes the core logic: standardized probes turn complex training into analyzable repeated micro-events without breaking immersion.

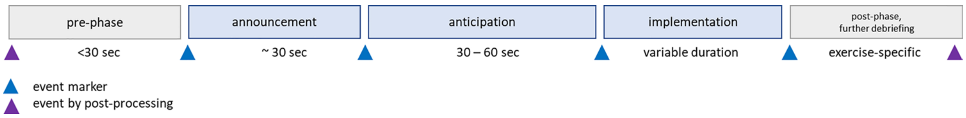


Figure 2: Stress-probe structure for comparable measurements in laboratory trials, field exercises and MR simulation. Short auditory-social probes are time-locked to biosignal windows, while longer stressor episodes support analysis of sustained stress and recovery (Paletta et al., 2025).

SYSTEM ARCHITECTURE AND SENSOR INTEGRATION

The SmartSkills architecture links four layers. The MR simulation layer presents a digital-twin environment enriched with virtual objects, role players and scenario logic. The sensor layer captures eye movements, pupil diameter, cardiovascular activity, temperature and optional movement or electrodermal signals. The human-factors layer estimates stress, cognitive load, fatigue and situation awareness. Finally, the decision-support layer visualizes critical episodes for trainers and supports after-action review. The instructor can in principle replay a training sequence, inspect probe-locked physiological responses and relate them to decisions, gaze behaviour and communication.

A key design principle is unobtrusive sensing. In the MR setting, the headset is combined with embedded eye tracking. Smart textiles and wearables provide cardiovascular data without restricting movement. In the field, mobile eye-tracking glasses collect gaze, video and audio for comparison with the simulated scenario. Figure 3 shows the system architecture (a) and core sensor components (b, c, d).

The sensor components are displayed as follows, Figure 3b shows the MR component with an embedding for a high precision, self-calibrating eye tracking component (Neon, Pupil Labs GmbH, Germany); Figure 3c denotes a smart textile cardiovascular measurement unit (Qus Tech GmbH, Austria) with temperature sensor (greenTEG AG, Switzerland), and Figure 3d refers to eye tracking glasses (video, audio; Pupil Labs GmbH, Germany) to collect data in the field at the real Native Challenge.

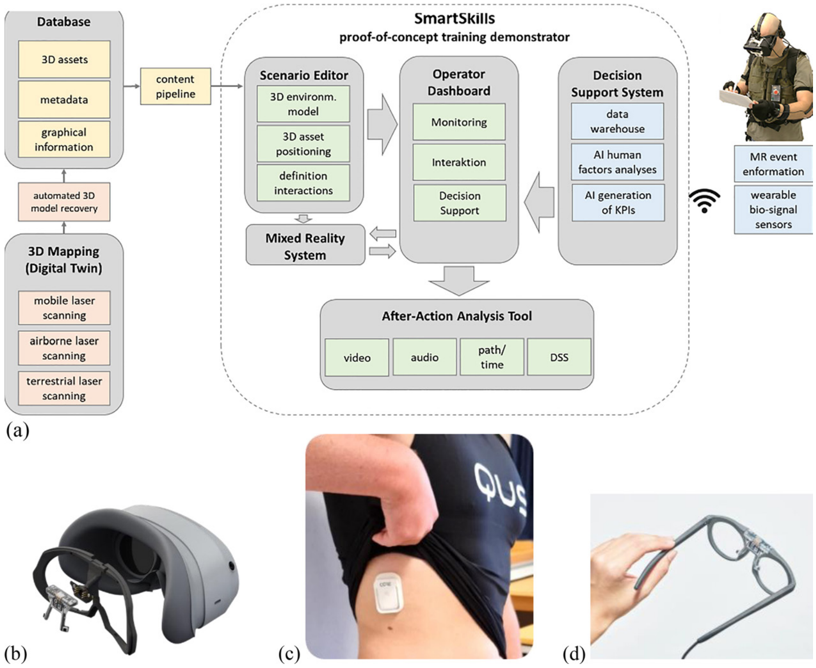


Figure 3: SmartSkills system architecture and sensing components: (a) digital twin, human-factors analytics, MR simulation and 3D visualization pipeline; (b) MR headset with eye-tracking embedding; (c) smart textile cardiovascular measurement unit; and (d) mobile eye-tracking glasses for field data collection.

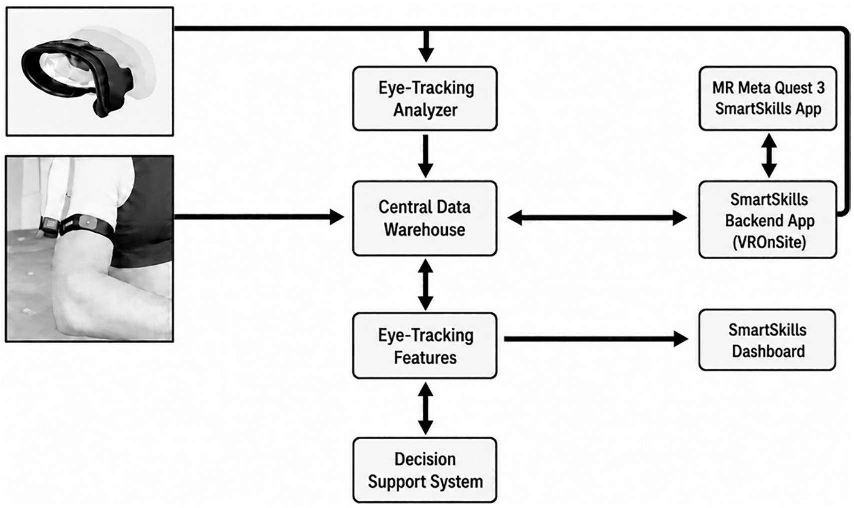


Figure 4: Biosignal-based sensor data (eye-tracking, cardiovascular) and MR-based data (Meta Quest 3) were streaming via Wi-Fi and BT to the central data warehouse (CDW). The CDW triggered expert dashboard and decision support system for trainers' insight and feedback.

Figure 4 summarizes the data-flow architecture used in the exploratory MR training setup. Eye-tracking data, cardiovascular sensor signals and MR interaction data from the Meta Quest 3 SmartSkills app are transmitted via Wi-Fi and Bluetooth into a Central Data Warehouse (CDW). The CDW serves as the synchronization and integration layer, linking raw sensor streams with eye-tracking analytics, derived features and scenario events. These processed indicators are then forwarded to the SmartSkills dashboard and the Decision Support System, enabling trainers to inspect stress-related responses, trainee state and relevant scenario episodes for feedback and debriefing.

The digital-twin pipeline acquires real environments using mobile laser scanning, drone-based mapping and vehicle- or backpack-based capture. Point clouds and images are fused, anonymized and transformed into textured 3D assets. Scenario designers can then place vehicles, obstacles, actors, medical objects and threat cues by drag-and-drop. For the checkpoint scenario, this allows rapid variation of visibility, escape options, actor positions and auditory event timing while keeping the training objective constant.

EXPLORATORY PILOT STUDY IN VIENNA 2025

The exploratory pilot study was conducted at the Johanniter Simulation Centre in Vienna in December 2025. The aim was proof-of-concept validation rather than hypothesis testing. The checkpoint use case and technical specification had been defined in advance, and the study focused on whether synchronized wearable and eye-tracking data could be captured during realistic MR checkpoint interactions.

Twelve participants took part (10 males, 2 females; age 37.82 +/- 14.39 years). Most participants reported no prior VR/MR/AR experience (10 of 12), and the educational level was predominantly upper-secondary or higher. The study comprised in total six trials. Each trial included one trainer and two trainees, indoor and outdoor components, a pre-session, baseline measurement, familiarization and a scenario session of approximately 5-10 minutes (Figure 5).

Figure 5: Implementation and debriefing context: (a) wearable biosignal sensor setup, (b) trainee in the MR simulation environment validating external calibration targets, (c) synchronized biosignal-based dashboard, and (d) mixed-reality training space (Johanniter, Vienna) with real actors and real-time displays with biosignal-based feedback from auditory stress-probes.

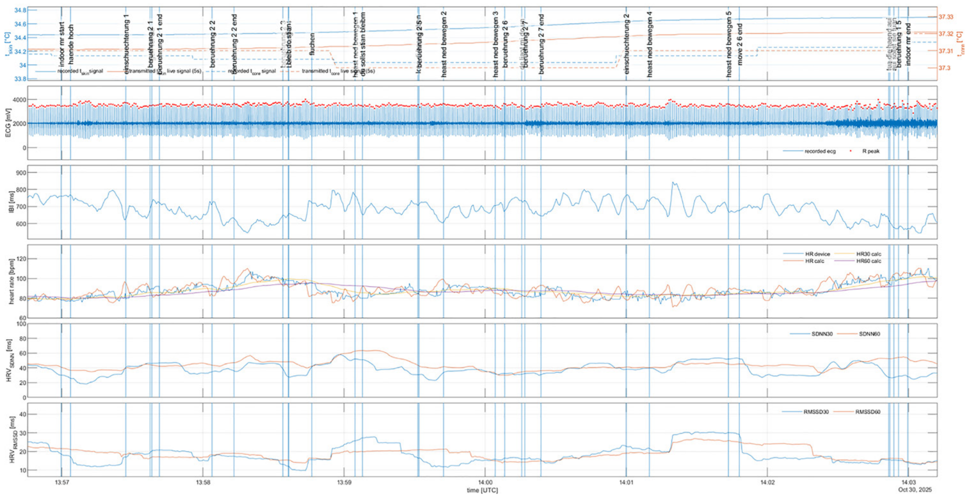


Figure 6: Synchronized pilot data from the Vienna 2025 proof-of-concept study. Event markers, wearable cardiovascular traces and eye-tracking information (not displayed) were aligned to support probe-based analysis and after-action review.

Measurement availability differed across trials; valid eye-tracking data were available for five sessions only, namely 1.2, 2.1, 4.1, 5.2, and 6.1, where the first digit denotes the trial number and the second digit identifies the corresponding participant. Short SP stress probes occurred three times in the context of pupillometric measurements: in each of sessions 1.2 and 2.1, once in session 5.2. ST stressors with longer shouting episodes were found in all sessions and were marked separately as sustained stressors with specific duration.

All data streams were synchronized around scenario events (Figure 6). The analysis pipeline aligned heart rate and RMSSD traces with stress-probe markers, extracted pupil-diameter time courses around auditory-social events and summarized both short probes and longer shouting episodes. Because of the pilot character, results are interpreted descriptively and used to refine synchronization, sensor placement, artefact handling and future statistical modelling.

EXPLORATORY RESULTS

The pilot confirmed that MR checkpoint training can generate analysable autonomic and pupillometric responses while preserving scenario realism.

In the cardiovascular data of participant 2.2 (Figure 7), short stress probes were associated with transient changes in heart rate and RMSSD that were observable in synchronized event windows. As expected for ultra-short windows, RMSSD was noisier and more sensitive to window length than heart rate. This supports using RMSSD primarily as a short-term recovery and parasympathetic modulation indicator, while treating single 2.3 second events with caution.

Pupillometry provided complementary temporal information. Eye-tracking data from session 2.1 showed clear event-related pupil dynamics around marked probes (Figure 8). The pupil response is especially useful because it captures orienting and arousal at sub-second resolution, but it also requires strict control for blink artefacts, gaze position and changing luminance in MR. The strongest value therefore lies in multimodal convergence: a probe is more informative when pupil dilation, gaze reallocation and cardiovascular change are temporally consistent.

Across the three eye-tracking sessions, the descriptive summary suggested that short probes and longer shouting episodes produce different temporal signatures. Short probes tend to evoke a phasic response, whereas sustained shouting produces broader activation and slower recovery. This distinction is important for training analytics: phasic reactivity can indicate orienting and threat detection, while sustained activation may index prolonged coping demands, emotion regulation load or scenario overload. Figure 9 summarizes the exploratory SP and ST curves across available sessions.

The Vienna 2025 pilot demonstrates the feasibility of combining MR, scenario event logging, wearable cardiovascular sensing and pupillometry in an illegal checkpoint training environment. The most important methodological gain is the transformation of an immersive and socially complex exercise into a set of analysable, repeated micro-events. Instead of reducing realism to laboratory stimuli, the stress probes preserve ecological validity by using scenario-congruent auditory and social cues. This is particularly relevant for checkpoint behaviour, where hostile voice, prosody, sudden commands and social dominance can trigger rapid physiological responses before any overt action occurs.

The approach also clarifies how biosignals should be used in training. Individual values should not be interpreted as simple pass-fail markers. Heart rate, RMSSD and pupil diameter are affected by movement, breathing, posture, luminance, display properties, task engagement and individual baseline differences. Their value is highest when they are synchronized with scenario events and interpreted together with behaviour: Did the trainee notice the threat? Did gaze move to the relevant actor? Was the team still communicating? Did physiological activation recover after de-escalation?

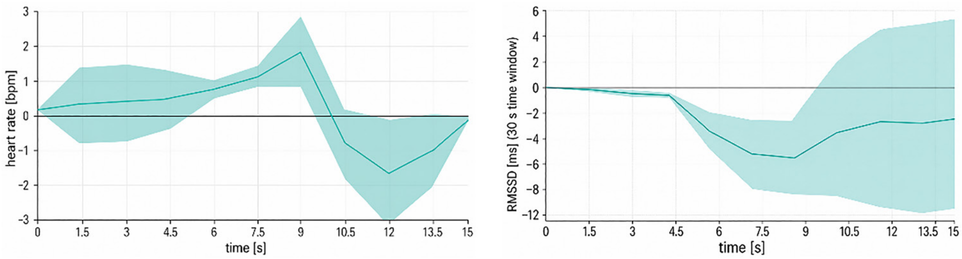


Figure 7: Cardiovascular (SP) probe-locked examples for participant 2.2: (a) heart-rate response across short stress probes and (b) RMSSD response across the same event windows. The plots illustrate the feasibility of aligning wearable cardiovascular features to short MR stressors.

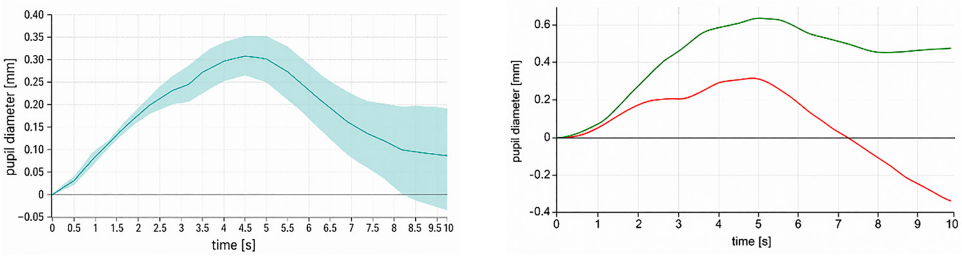


Figure 8: Pupillometric (SP) probe responses from eye-tracking sessions: (a) all available curves from session 2.1 and (b) single responses to the first (green) and second (red) appearance of the stimulus in a serious of stress-probes. The curves support event-related interpretation of short auditory-social stressors.

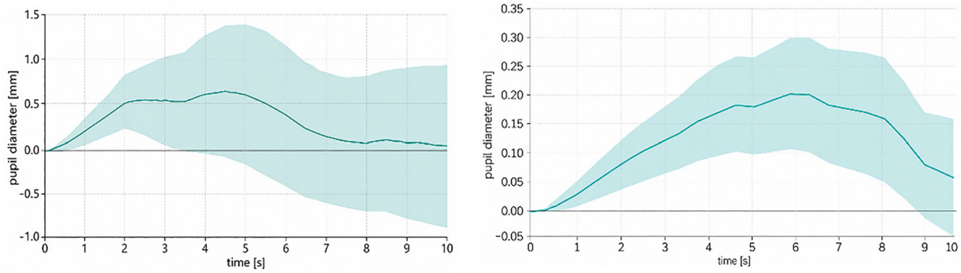


Figure 9: Summary of eye-tracking curves across available sessions: (a) short stress-probe responses and (b) longer shouting-stressor responses. The difference between phasic and sustained response shapes is central for future adaptive training logic.

For trainers, the resulting decision support should therefore emphasize episodes, not diagnoses. A useful dashboard flags moments where physiological activation, gaze behaviour and task performance diverge from expected patterns. Instructors can then replay the episode and discuss communication, positioning, attention and emotional regulation with the trainee. In later development stages, the same information can support adaptive scenario difficulty, for example by increasing actor hostility only when a trainee shows stable performace and manageable stress response.

The Vienna 2025 pilot therefore operationalized stress probes as short, standardized, time-locked events embedded into the illegal checkpoint scenario. Conceptually, these probes combine principles from laboratory stress induction (Kirschbaum et al., 1993), social-evaluative threat and uncontrollability theory (Dickerson & Kemeny, 2004), and auditory arousal research showing pupil responses to emotionally salient sounds (Widmann et al., 2018). The synchronized assessment of pupil dynamics, HR and RMSSD follows the rationale of multimodal autonomic arousal profiling (Wang et al., 2018).

LIMITATIONS AND FUTURE WORK

The study is exploratory, with a small sample, heterogeneous measurement availability and limited control over environmental factors. It was not designed to test group differences or establish diagnostic thresholds. Ultra-short HRV windows require particular caution, and pupillometry in MR must control luminance, gaze position and blink-related artefacts. Future work will extend the sample, standardize the probe library, include EDA and respiration, compare MR with field exercises, and test whether probe-based indicators predict trainer ratings, team performance and retention of de-escalation skills.

The next methodological step is a preregistered validation study with predefined feature windows, scenario variants and mixed-effects models that separate person, trial, probe type and environment effects. A second step is closed-loop adaptation: the system should not merely measure stress but use validated indicators to personalize training intensity while keeping the instructor in control.

CONCLUSIONS

SmartSkills advances MR training from immersive visualization toward data-driven human-factors learning analytics. By embedding short auditory-social stress probes and longer shouting episodes into a realistic illegal checkpoint scenario, the system creates comparable physiological measurement windows without sacrificing ecological validity. The Vienna 2025 pilot shows that synchronized heart-rate, RMSSD and pupillometry data can be captured during realistic MR training and can support after-action review. The resulting framework provides a promising basis for adaptive stress-resilience training, objective debriefing and future digital-twin based simulation centres for peace missions and first responders.

ACKNOWLEDGMENT

This work was partially funded by the projects SmartSkills (grant No. FO999905346), EMERDEC (grant No. FO999914127), and SmartResponse (grant No. FO999926240) of the Austrian FFG KIRAS programme, co-financed by the Austrian Federal Ministry for Innovation, Mobility and Infrastructure (BMIMI).

REFERENCES

Arnal, L. H., Flinker, A., Kleinschmidt, A., Giraud, A.-L., and Poeppel, D. (2015). Human screams occupy a privileged niche in the communication soundscape. Current Biology, 25(15), 2051-2056. https://doi.org/10.1016/j.cub.2015.06.043.

Beaurenaut, M., Tokarski, E., Dezecache, G., and Grèzes, J. (2020). The Threat of Scream paradigm: a tool for studying sustained physiological and subjective anxiety. Scientific Reports, 10, 12496. https://doi.org/10.1038/s41598-020-68889-0.

Bradley, M. M., and Lang, P. J. (2000). Affective reactions to acoustic stimuli. Psychophysiology, 37(2), 204-215. https://doi.org/10.1111/1469-8986.3720204.

Bradley, M. M., Miccoli, L., Escrig, M. A., and Lang, P. J. (2008). The pupil as a measure of emotional arousal and autonomic activation. Psychophysiology, 45(4), 602–607. https://doi.org/10.1111/j.1469-8986.2008.00654.x.

Castaldo, R., Melillo, P., Bracale, U., Caserta, M., Triassi, M., and Pecchia, L. (2015). Acute mental stress assessment via short term HRV analysis in healthy adults: a systematic review with meta-analysis. Biomedical Signal Processing and Control, 18, 370–377. https://doi.org/10.1016/j.bspc.2015.02.007.

Chen, K.-H., Aksan, N., Anderson, S. W., Grafft, A. and Chapleau, M.W. (2014) Habituation of parasympathetic-mediated heart rate responses to recurring acoustic startle. Front. Psychol. 5:1288. https://doi.org/10.3389/fpsyg.2014.01288.

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.

Driskell, J. E., Johnston, J. H., & Salas, E. (2001). Does stress training generalize to novel settings? Human Factors, 43(1), 99–110. https://doi.org/10.1518/001872001775992471

Enders, L. R., Gordon, S. M., Roy, H., Rohaly, T., Dalangin, B., Jeter, A., Villarreal, J., Boykin, G. L., and Touryan, J. (2024). Evidence of elevated situational awareness for active duty soldiers during navigation of a virtual environment. PLoS ONE, 19(5), e0298867. https://doi.org/10.1371/journal.pone.0298867.

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.

Farra, S. L., Miller, E. T., Timm, N., and Schafer, J. (2013). Improved training for disasters using 3-D virtual reality simulation. Western Journal of Nursing Research, 35(5), 655–671. https://doi.org/10.1177/0193945912471735.

Giannakakis, G., Grigoriadi, D., Giannakaki, K., Simantiraki, O., Roniotis, A., and Tsiknakis, M. (2022). Review on psychological stress detection using biosignals. IEEE Transactions on Affective Computing, 13(1), 440–460. https://doi.org/10.1109/TAFFC.2019.2927337.

Gomez, P., and Danuser, B. (2004). Affective and physiological responses to environmental noises and music. International Journal of Psychophysiology, 53(2), 91–103. https://doi.org/10.1016/j.ijpsycho.2004.02.002.

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. International Journal of Human-Computer Studies, 66(9), 641–661. https://doi.org/10.1016/j.ijhcs.2008.04.004.

Kallus, K. W., & Gaisbachgrabner, K. (2018). Stress and recovery in applied settings: Long working hours, recovery, and breaks. In M. Kellmann & J. Beckmann (Eds.), Sport, recovery, and performance: Interdisciplinary insights (pp. 233–246). Routledge. https://doi.org/10.4324/9781315268149-16

Kim, H.-G., Cheon, E.-J., Bai, D. S., Lee, Y. H., and Koo, B.-H. (2018). Stress and heart rate variability: a meta-analysis and review of the literature. Psychiatry Investigation, 15(3), 235–245. https://doi.org/10.30773/pi.2017.08.17.

Kirschbaum, C., Pirke, K. M., and Hellhammer, D. H. (1993). The Trier Social Stress Test: a tool for investigating psychobiological stress responses in a laboratory setting. Neuropsychobiology, 28(1–2), 76–81. https://doi.org/10.1159/000119004.

Laborde, S., Mosley, E., and Thayer, J. F. (2017). Heart rate variability and cardiac vagal tone in psychophysiological research: recommendations for experiment planning, data analysis, and data reporting. Frontiers in Psychology, 8, 213. https://doi.org/10.3389/fpsyg.2017.00213.

Montes-Lourido, P., Kar, M., Kumbam, I. et al. (2021) Pupillometry as a reliable metric of auditory detection and discrimination across diverse stimulus paradigms in animal models. Sci Rep 11, 3108. https://doi.org/10.1038/s41598-021-82340-y

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., Pszeida, M., Schneeberger, M., Dini, A., Reim, L., and Kallus, K. W. (2022). Cognitive-emotional stress and risk stratification of situational awareness in immersive first responder training. IEEE-EMBS International Conference on Biomedical and Health Informatics (BHI), Ioannina, Greece, 1-4.

Paletta, L., Peer, A., Öttl, M., Aumayr, G., Kallus, K., et al. (2025). A Framework for Mixed Reality-supported Training of Conflict Resolution and First Responder Skills in International Crisis Situations: SmartSkills. In: Lucas Paletta (eds) Cognitive Computing and Internet of Things. AHFE (2025) International Conference. AHFE Open Access, vol 165. AHFE International, USA. http://doi.org/10.54941/ahfe1006097

Pan, J., et al. (2024). The effects of emotional arousal on pupil size depend on luminance. Scientific Reports, 14, 21164. https://doi.org/10.1038/s41598-024-70895-5.

Partala, T., and Surakka, V. (2003). Pupil size variation as an indication of affective processing. International Journal of Human-Computer Studies, 59(1-2), 185–198. https://doi.org/10.1016/S1071-5819(03)00017-X.

Salmon, P. M., Stanton, N. A., Walker, G. H., et al. (2009). Measuring situation awareness in complex systems: comparison of measures study. International Journal of Industrial Ergonomics, 39(3), 490–500. https://doi.org/10.1016/j.ergon.2008.10.010.

Saunders, T., Driskell, J. E., Johnston, J. H., & Salas, E. (1996). The effect of stress inoculation training on anxiety and performance. Journal of Occupational Health Psychology, 1(2), 170–186. https://doi.org/10.1037/1076-8998.1.2.170

Schneeberger, M., Paletta, L., Kallus, K.W., et al. (2022). First Responder Situation Reporting in Virtual Reality Training with Evaluation of Cognitive-emotional Stress using Psychophysiological Measures, In: Paletta, L., Ayaz H., eds., Cognitive Computing and Internet of Things, AHFE international open access, http://doi.org/10.54941/ahfe1001841

Setz, C., Arnrich, B., Schumm, J., La Marca, R., Tröster, G., and Ehlert, U. (2010). Discriminating stress from cognitive load using a wearable EDA device. IEEE Transact on Inf Techn in Biomed, 14(2), 410–417. https://doi.org/10.1109/TITB.2009.2036164.

Shaffer, F., and Ginsberg, J. P. (2017). An overview of heart rate variability metrics and norms. Frontiers in Public Health, 5, 258. https://doi.org/10.3389/fpubh.2017.00258.

Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology. (1996). Heart rate variability: standards of measurement, physiological interpretation and clinical use. Circulation, 93(5), 1043–1065. https://doi.org/10.1161/01.CIR.93.5.1043.

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.

Zekveld, A. A., Koelewijn, T., and Kramer, S. E. (2018). The pupil dilation response to auditory stimuli: current state of knowledge. Trends in Hearing, 22, 1–25. https://doi.org/10.1177/2331216518777174.