When Simulation Meets Physiology: Adaptive Resilience for Mission-Critical XR in Fast-Jet Training

Julia Schorlemmer[1,2,*], Alia Saad[1], Sebastian Möller[2,3] and Jan-Niklas Voigt-Antons[1]

[1]*University of Applied Sciences Hamm-Lippstadt, Marker Allee 76–78, 59063 Hamm, Germany*
[2]*Technische Universität Berlin, Straße des 17. Juni 135, 10623 Berlin, Germany*
[3]*German Research Center for Artificial Intelligence (DFKI), Alt-Moabit 91c, 10559 Berlin, Germany*

Abstract

Extended Reality (XR) technologies are increasingly explored for training in safety-critical domains. However, many XR design paradigms implicitly assume that seamlessness, immersion, and automation are inherently beneficial. This assumption is not necessarily applicable in high-stakes training contexts where simulated environments intersect with physiological limits. This paper examines fast-jet pilot training as an extreme but revealing case of mission-critical XR. We show how inadequate physiological sensing, opaque adaptation, or misleading feedback can silently undermine training validity rather than enhance performance. To mitigate these risks, we introduce an adaptive resilience layer for mission-critical XR training systems, prioritizing physiology-aware calibration, transparent uncertainty signaling, conservative adaptation, and situated feedback. These principles are essential for safeguarding the validity, calibrated trust, and human agency under stress. Although illustrated through XR-based g-force training for the Anti-G Straining Maneuver (AGSM), the insights derived from this case extend beyond simulation-based training and are relevant to XR systems that support professionals in both training and live operational contexts.

Keywords

XR training, mission-critical XR, fast-jet aviation, physiology-aware adaptation, adaptive resilience

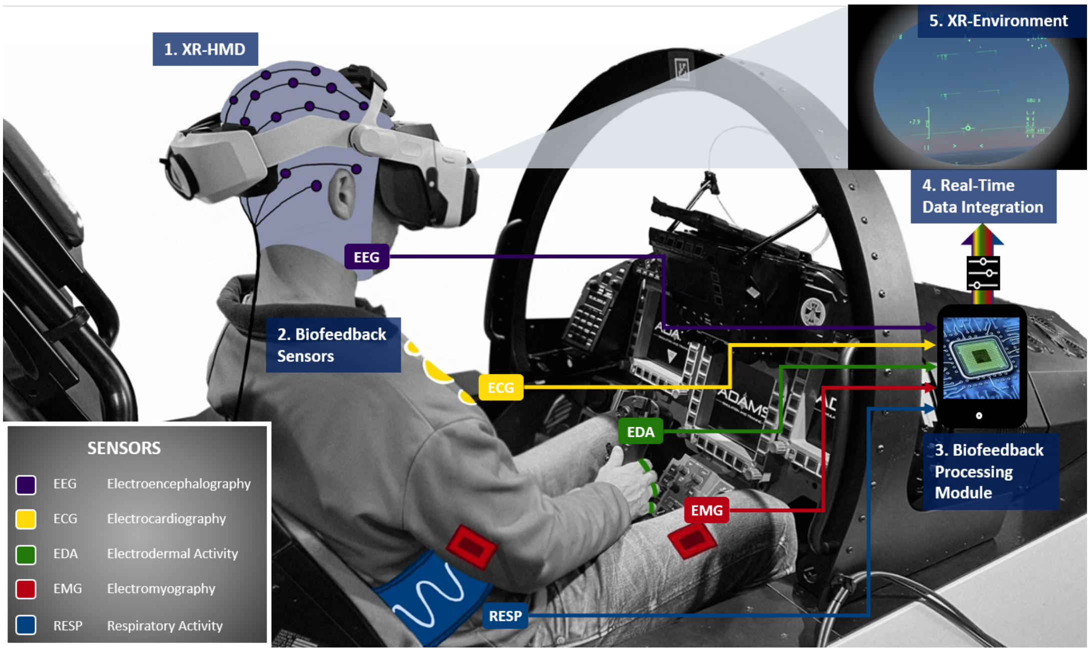

Figure 1: Conceptual illustration of adaptive resilience principles for mission-critical XR training environments. The framework highlights how physiological sensing, uncertainty-aware feedback, and conservative adaptation mechanisms help maintain training validity under extreme conditions such as fast-jet pilot training (illustration provided by Adams Simulation and Training GmbH, used with permission).

CHI'26: XR4CE workshop, April 14, 2026, Barcelona, Spain
*Corresponding author.
julia.schorlemmer@hshl.de (J. Schorlemmer); alia.saad@hshl.de (A. Saad); sebastian.moeller@tu-berlin.de (S. Möller); jan-niklas.voigt-antons@hshl.de (J. Voigt-Antons)

1. Fast-Jet Training as a Challenging Environment

Fast-jet pilot training constitutes a particularly demanding form of high-stakes, embodied work that differs fundamentally from many commonly studied domains of Extended Reality (XR). Beyond training scenarios, fast-jet aviation in general illustrates a broader class of mission-critical XR contexts in which immersive technologies may support professionals during live operational work, where system reliability, human performance, and safety are tightly coupled. In real flight, pilots operate under extreme physical loads, characterized by rapid onset rates of positive g-acceleration that can exceed physiological tolerance limits within seconds [1]. At the same time, pilots are required to maintain high levels of cognitive performance under severe time pressure, managing complex mission tasks, system monitoring, and decision-making in parallel. Errors in this context are not merely performance degradations but can have irreversible and catastrophic consequences, including loss of aircraft control or loss of consciousness [1, 2, 3].

As a result, training quality in fast-jet contexts is inherently safety-critical. Training systems not only shape task performance but also directly influence the development and maintenance of reflexive, embodied skills that are essential for flight safety [4]. In contrast to low-stakes or exploratory training scenarios, inadequate or misleading training feedback can undermine critical physiological responses and foster unsafe mental models. Consequently, the primary objective of training systems in this context is not maximal performance or immersion, but the preservation of training validity under conditions of extreme physiological and cognitive stress - that is, the extent to which training reliably elicits realistic cognitive, physiological, and behavioral responses that transfer to real flight conditions.

These characteristics impose non-negotiable requirements on XR-based training systems. First, **individual physiological calibration** is essential, as physiological tolerance, cognitive capacity, and stress responses vary substantially between pilots and across training states. Second, **system breakdowns and uncertainty** must be handled explicitly rather than concealed; silent failures in sensing, inference, or timing risk producing invalid training conditions without the trainee's awareness. Third, **explanations and feedback** must remain interpretable under high cognitive and physical load, operating at a glance and without imposing additional attentional demands.

Conventional XR design goals, such as seamlessness, maximal immersion, or increased levels of automation, are therefore insufficient in this setting. While such goals may enhance user experience in other domains, they can become problematic in safety-critical training environments by obscuring system limits, masking degradation, and fostering misplaced trust. Therefore, XR-based training approaches for fast-jet pilots must be evaluated not only in terms of usability and immersion, but also in terms of whether they maintain training validity under extreme physiological and cognitive stress. In the following, we introduce XR-based g-force training for the Anti-G Straining Maneuver (AGSM) as a concrete, safety-critical use case to ground this discussion.

2. XR-Based G-Force and AGSM Training as a Safety-Critical Use Case

Exposure to high positive g-loads poses a well-established risk of almost loss of consciousness (a-LOC) or g-induced loss of consciousness (g-LOC), directly threatening flight safety [5, 6]. Preventing these events depends on the timely and sufficiently strong execution of the AGSM, a coordinated pattern of muscular tension and respiratory control. Once visual impairment or cognitive degradation begins, corrective actions are often too late, making early reflexive responses more critical than conscious control. These physiological limits encountered during high-g flight pose fundamental challenges for mission-critical XR systems designed to support professionals operating under extreme physical and cognitive stress, whether in simulation-based training or potential operational assistance contexts. In real flight operations, military pilots already use helmet-mounted displays integrated into their flight helmets. These systems enhance situational awareness by projecting mission-critical information directly into the pilot's field of view, overlaying digital flight data onto the external environment in real time. However, current helmet-mounted display systems do not incorporate real-time physiological sensing capable of monitoring or supporting the pilot's physiological state during high-g exposure. Also, in the context of AGSM training, supporting viable functions of the pilot with XR systems, aims not at procedural knowledge alone, but at stabilizing an embodied reflex that must be triggered automatically under stress [6].

Historically, AGSM-related reflexes have not been trained through simulators. While simulators are central to fast-jet training, maintaining AGSM effectiveness relies on real-flight exposure and physiological training, as the relevant training signal is the pilot's bodily response to sustained g-load rather than the maneuver itself. This marks a structural boundary of simulation: while perceptual correlates of g-exposure (e.g., visual impairment or warning cues) may be approximated, the cardiovascular and muscular strain that gives rise to AGSM reflexes cannot be generated by XR. XR-based training, therefore, cannot replace physiological exposure, but must be designed to support and preserve reflexes acquired through real-world training. This raises the question of how XR systems may incorporate physiological sensing in order to monitor and support these embodied responses during simulated training.

To date, XR-based training systems for fast-jet pilots have not systematically attempted to simulate g-loads in order to preserve or reinforce AGSM reflex timing through controlled visual, haptic, or adaptive feedback. As a result, when simulated flight scenarios are experienced without corresponding physiological symptoms, pilots may enforce and tolerate g-profiles that would be unsustainable in real flight. This could potentially decouple perceived task demands from required physical effort and timing, creating a critical design risk for XR-based training systems.

Addressing this boundary case is therefore not merely hypothetical but forms the basis of ongoing conceptual and technical exploration into how XR systems might support, rather than inadvertently weaken, embodied safety-critical reflexes under conditions of sensing uncertainty and limited physiological coupling.

A first prototype integrating multimodal physiological sensing as well as haptic, visual, and adaptive XR feedback is currently under development in collaboration with an industry partner specializing in fast-jet simulation systems. The system combines physiological monitoring (e.g., electromyography, respiration, heart rate variability, electrodermal activity and electroencephalography) with adaptive XR simulation control in order to explore how individualized calibration, uncertainty-aware feedback, and conservative adaptation strategies can be implemented in practice. Similar approaches have been explored in physiologically-adaptive MR systems that use biosignals as inputs to closed-loop adaptive interfaces [7].

The g-force and AGSM use case exposes a fundamental asymmetry of XR training: perceptual realism can be approximated, but physiological equivalence cannot. Designing explicitly for this asymmetry, rather than concealing it, is essential for preserving training validity in safety-critical contexts. A related asymmetry is also present in real-flight environments. Fast-jet pilots already receive digital mission-critical information through helmet-mounted displays integrated into their flight helmets, yet these systems currently operate without real-time physiological sensing. Whenever immersive digital systems intersect with human physiological limits, physiology-aware system design becomes essential and must explicitly account for incomplete sensing, uncertain inference, and the need to maintain calibrated human trust across both training environments and potential operational XR support systems.

3. When XR Training Fails: Risks and Consequences in High-Stakes Contexts

XR technologies offer substantial potential for training in challenging environments and for supporting professionals in safety-critical operational contexts [8]. In safety-critical contexts such as XR-based g-force and AGSM training, however, system behavior becomes part of the training signal itself, magnifying the impact of design decisions beyond usability.

3.1. Sources of Risk in XR-Based Training

In the context of XR-based g-force training, several sources of risk become particularly salient. **Inter- and intra-individual variability** poses a fundamental challenge, as physiological tolerance, cognitive capacity, and stress responses differ substantially between pilots and fluctuate across training states. Relying on population-level assumptions or generic thresholds, therefore, risks producing misleading feedback in scenarios that critically depend on individualized reflex timing and physical effort. At the same time, **sensor-related uncertainty** represents an inherent limitation of XR systems that integrate physiological and behavioral sensing. Noise, latency, tracking loss, and synchronization errors are especially likely under movement-intensive, physiologically demanding conditions, such as AGSM execution, directly affecting the reliability of system inference and feedback timing [9]. In addition, **system-induced cognitive load** can emerge when interaction is unstable, feedback is delayed, or visualizations are overly complex. Such extraneous load competes with time-critical perception and action during high-stress training tasks, potentially interfering with learning and performance rather than supporting them [10]. Finally, **misplaced design priorities** can exacerbate these risks. Emphasizing realism, immersion, or seamless adaptation over training validity may obscure system limitations and divert attention from whether the XR system meaningfully supports the acquisition of safety-critical, transferable skills.

3.2. Harmful Consequences for Training Validity

If these risks are not explicitly addressed, XR-based training could produce harmful effects, particularly in safety-critical use cases. **Negative training effects** may occur when reflexive skills are weakened, mistimed, or insufficiently trained, fostering unsafe assumptions about physiological tolerance or task demands that do not transfer to real-world conditions. At the same time, **loss of calibrated trust** can arise from inconsistent sensing, delayed feedback, or opaque system behavior. Under such conditions, trainees may struggle to assess whether system feedback is reliable in a given situation, increasing the risk of over- or under-reliance on the training system. Furthermore, **cognitive interference with mission-critical processes** can result when system-induced load distorts attention, decision-making, or learning priorities during time-critical training phases. Rather than supporting performance, the XR system may compete with the very cognitive resources required for effective action under stress. Finally, **erosion of self-regulation and agency** can emerge when adaptive systems compensate silently for inadequate execution or signal success despite poor technique. Such behavior prevents trainees from recognizing deficits, undermining self-assessment, self-correction, and the long-term development of transferable skills.

3.3. Implications for XR Training Systems

Taken together, these risks and consequences suggest that physiology-aware XR systems for safety-critical environments cannot be designed to maximize seamlessness, immersion, or automation alone. Prior research on safety-critical automation highlights that high levels of automation can obscure system limits and impact operator trust unless uncertainty and reliability are explicitly communicated [11, 12]. Similarly, XR reviews in high-risk domains note that systems that ignore the complexity of physiological and cognitive stress may create conditions that appear valid in simulation but fail to transfer safely to real environments [8]. In the context of the g-force and AGSM training use case, this reframes the design objective: rather than optimizing system performance alone, XR systems must safeguard the acquisition of valid, transferable skills under conditions of physiological and cognitive stress.

We therefore propose an *adaptive resilience layer* as a guiding design stance for mission-critical XR systems. Rather than specifying a fixed system architecture, this layer articulates design principles for maintaining user agency, calibrated trust in the system, and individualized support under conditions where sensing becomes unreliable, feedback timing is uncertain, or physiological coupling remains incomplete. Importantly, these principles extend beyond training environments. In operational flight contexts, pilots cannot rely on the protective structure of a training setting but must apply learned skills and reflexes under real mission conditions. In such environments, system transparency, conservative adaptation, and calibrated trust become even more critical for safe human–machine collaboration.

At a minimum, such a resilience-oriented design stance implies four tightly coupled properties:

- **(1) physiology-aware calibration** to account for inter- and intra-individual variability by grounding feedback and adaptation in individualized baselines rather than population-level assumptions; for example, calibrating respiratory or muscular strain thresholds individually before XR-based g-force training, or adapting decision-support thresholds in operational XR systems based on an operator's baseline physiological workload.
- **(2) self-diagnostics and explicit uncertainty signaling** to make degradation and breakdowns visible instead of concealing them, thereby preventing misplaced trust under invalid system states; for instance, indicating degraded physiological sensing through subtle peripheral indicators in a head-mounted display or cockpit helmet-mounted display, allowing pilots to immediately recognize when system inference may be unreliable.
- **(3) graceful degradation through conservative adaptation** rather than silent compensation, ensuring that the system does not artificially stabilize performance or conceal insufficient human effort, thereby preserving self-regulation and agency; for example, temporarily suspending adaptive feedback when sensor quality degrades, reverting to conservative task parameters during training, or reducing automated assistance in operational XR systems when uncertainty exceeds safe thresholds. In such cases, simple rule-based guidance may override adaptive behavior in order to prioritize safety during critical situations.
- **(4) situated, low-load feedback** that supports learning and self-regulation while minimizing extraneous cognitive load, particularly under conditions of high cognitive and physiological stress. In practice, this may involve peripheral visual cues, simple auditory tones, or minimal haptic seat feedback that communicate system state without requiring focal attention, allowing users to remain focused on mission-critical tasks.

Figure 1 illustrates how these considerations emerge in g-force training contexts, where tightly coupled physiological sensing (e.g., EMG, respiration, ECG, EDA, EEG) and XR feedback amplify the consequences of sensing uncertainty and design decisions. The g-force use case thus serves as a boundary case that highlights why resilience, calibration, and situated explainability should be considered foundational requirements for physiology-aware XR systems operating in high-stakes contexts.

While this paper focuses on XR-based fast-jet training as an illustrative boundary case, the proposed resilience layer extends to a broader class of safety-critical XR applications. Whenever immersive systems intersect with human physiological limits, design must account for sensing uncertainty, cognitive stress, and time-critical decision-making. Across domains such as aviation decision support, emergency response, or medical procedures, XR systems should therefore be designed for resilience rather than seamlessness - a requirement whenever immersive technologies operate at the boundary of human physiological limits.

Acknowledgments

The authors thank Adams Simulation and Training GmbH for their collaboration and valuable conceptual input during the development of the XR-based fast-jet training use case presented in this paper. Illustrative material and domain expertise were provided as part of this collaboration.

Declaration on Generative AI

During the preparation of this work, the authors used ChatGPT and Grammarly in order to check grammar and spelling, paraphrase, and reword existing text. After using these tools, the authors reviewed and edited the content as needed and take full responsibility for the publication's content.

References

[1] M.-Y. Tu, H. Chu, Y.-J. Lin, K.-T. Chiang, C.-M. Chen, H.-H. Chen, C.-S. Yang, C.-Y. Lai, Combined effect of heart rate responses and the anti-g straining manoeuvre effectiveness on g tolerance in a human centrifuge, Scientific Reports 10 (2020) 21611. doi:10.1038/s41598-020-78687-3.

[2] A. Hamann, N. Carstengerdes, Assessing the development of mental fatigue during simulated flights with concurrent eeg-fnirs measurement, Scientific Reports 13 (2023). doi:10.1038/s41598-023-31264-w.

[3] A. Landman, E. L. Groen, M. M. R. van Paassen, A. W. Bronkhorst, M. Mulder, Dealing with unexpected events on the flight deck: a conceptual model of startle and surprise, Human Factors 59 (2017) 1161–1172. doi:10.1177/0018720817723428.

[4] T. Marron, N. Dungan, B. Mac Namee, A. D. O'Hagan, Virtual reality & pilot training: existing technologies, challenges & opportunities, Journal of Aviation/Aerospace Education & Research 33 (2024). URL: https://commons.erau.edu/jaaer/vol33/iss1/1/.

[5] S. Kim, T. Cho, Y. Lee, H. Koo, B. Choi, D. Kim, G-loc warning algorithms based on emg features of the gastrocnemius muscle, Aerospace Medicine and Human Performance 88 (2017) 737–742. doi:10.3357/AMHP.4781.2017.

[6] L. Yang, M. Dai, S. Li, H. Wang, X. Cao, Z. Zhao, Real-time assessment of global and regional lung ventilation in the anti-gravity straining maneuver using electrical impedance tomography, Computers in Biology and Medicine 135 (2021). doi:10.1016/j.compbiomed.2021.104592.

[7] F. Chiossi, S. Mayer, How can mixed reality benefit from physiologically-adaptive systems? challenges and opportunities for human factors applications, arXiv preprint arXiv:2303.17978 (2023). URL: https://arxiv.org/abs/2303.17978. doi:10.48550/arXiv.2303.17978.

[8] J. E. Dodoo, H. Al-Samarraie, A. I. Alzahrani, T. Tang, Xr and workers' safety in high-risk industries: a comprehensive review, Safety Science 185 (2025). doi:10.1016/j.ssci.2025.106804.

[9] X. Li, W. He, P. O. Kristensson, How do we evaluate experiences in immersive environments?, in: Proceedings of the CHI Conference on Human Factors in Computing Systems, Association for Computing Machinery, New York, NY, USA, 2026. URL: https://arxiv.org/abs/2601.17811.

[10] K. M. Stanney, J. Archer, A. Skinner, C. Horner, C. Hughes, N. P. Brawand, E. Martin, S. Sanchez, L. Moralez, C. M. Fidopiastis, R. S. Perez, Performance gains from adaptive extended reality training fueled by artificial intelligence, The Journal of Defense Modeling and Simulation 19 (2022) 195–218. doi:10.1177/15485129211064809.

[11] P. A. Hancock, R. J. Jagacinski, R. Parasuraman, C. D. Wickens, G. F. Wilson, D. B. Kaber, Human-automation interaction research: past, present, and future, Ergonomics in Design 21 (2013) 9–14. doi:10.1177/1064804613477099.

[12] S. C. Kohn, E. J. de Visser, E. Wiese, Y.-C. Lee, T. H. Shaw, Measurement of trust in automation: A narrative review and reference guide, Frontiers in Psychology 12 (2021). doi:10.3389/fpsyg.2021.604977.