I Trust You Body: Using Augmented Avatars in Challenging Environments

Clara Sayffaerth[1]

[1]LMU Munich, Human-Centered Ubiquitous Media, Frauenlobstraße 7a, Munich, 80337, Germany

Abstract

Extended Reality (XR) avatar guidance in Challenging Environments (CEs) promises to improve aspects such as performance, co-presence, and trust, but carries the risk of compromising human safety, situational awareness, and the ability to act under stress. This problem is critical when minor fluctuations in attention can lead to failures or even serious injuries. In this paper, we investigate the design of avatar-based guidance in XR systems for CEs. Based on previous work, we describe how avatars representing humans, AI, or hybrid solutions can be used in these contexts and identify future research directions. We argue that avatars should be understood not as static assistants, but as situational, adaptive guides that balance instructions and environmental awareness. Moreover, we see avatars as key mechanisms for calibrating emotional states, performance, and situational understanding in future XR systems for high-risk environments.

Keywords

Extended Reality, Avatars, Guidance, Social Interaction, Challenging Environments

Figure 1: Avatars can help handle stressful or dangerous environments by adapting the visualization based on the context to keep users safe, calm, and attentive.

1. Introduction

Recent advances in Extended Reality (XR) and Artificial Intelligence (AI) have expanded the design space for interactive applications. XR systems can now go beyond static overlays or pre-authored instructions to provide context-aware, real-time guidance due to AI-driven perception, reasoning, and adaptation. These capabilities enable virtual objects to be dynamically placed, modified, and removed based on sensor data or user behavior. Such systems open up new ways in complex, unpredictable, and dangerous real-world scenarios. These Challenging Environments (CEs) are, for example, factories, construction sites, mines, or disaster-response settings, where users operate heavy machinery, navigate unstable terrain, or face rapidly changing conditions.

One promising approach in this space is the use of avatars as guides or collaborators, providing step-by-step instructions, feedback, or proxies that are spatially anchored in the environment [1, 2, 3]. Such avatars can be mediators for real humans, autonomous AI agents, or hybrid approaches based on

CHI'26: XR4CE Workshop, April 14, 2026, Barcelona, Spain

clara.sayffaerth@ifi.lmu.de (C. Sayffaerth)

the use case [4]. Compared to abstract visual or textual cues, avatars offer a more embodied and social form of interaction, enabling users to interpret information through familiar human-like behaviors.

Prior research highlights the effectiveness of avatar-based instruction. Avatars have been shown to improve task performance in augmented environments, particularly during two-handed or spatial tasks [1], while also enhancing perceived co-presence and trust [5]. These benefits are especially relevant in time-critical situations that require rapid comprehension of complex procedures and precise coordination of actions. However, the design and visualization of avatars also have a substantial impact on user experience and behavior. Previous studies demonstrate that avatar characteristics can influence how users feel [6, 7], behave [8, 9, 10], perform [11], and remember instructions [12].

In particular, design factors such as realism, scale, perspective, and spatial positioning can potentially affect cognitive load, attention, and risk perception. For instance, a large or highly detailed avatar may increase engagement and strengthen social cues, but can also divert attention from the physical environment [13]. Virtual elements may occlude critical real-world information or draw attention away from potential hazards [14]. Such occlusion and attentional tunneling effects [15] are especially problematic in safety-critical contexts, where even brief visual obstruction or misaligned attention can have severe consequences. As a result, avatar design introduces inherent trade-offs between performance, user experience, and physical safety.

Given these challenges, the design of instructional XR systems for CEs must carefully balance the benefits of guidance with the need to maintain continuous awareness of the surrounding environment. Determining how much information to present, when to present it, and in what form remains an open research question, particularly when avatars are involved. Achieving this balance is critical to ensuring that XR systems support users effectively without introducing additional safety risks.

In this paper, we examine the opportunities and challenges of avatar-based instruction in CEs. We focus on how avatar design and visualization influence user performance, experience, situational awareness, safety, and derive implications for the design of future XR systems.

2. Using Avatars for Instructions

XR avatars have been widely explored to provide guidance and education, often in combination with AI-driven perception and adaptation [4, 16]. Prior research has investigated avatar-based instruction across a range of tasks and visualization strategies, frequently varying avatar appearance, embodiment, and perspective to assess their influence on user performance and experience. While avatars are often used in synchronous settings to mediate between remote users through telepresence [17], more research is analyzing asynchronous use cases where recordings or AI agents provide instructions [3, 2]. These studies demonstrate that avatar-based guidance is beneficial, especially for tasks involving complex, body-coordinated movements, where spatial alignment and motion timing are critical [1]. Sayffaerth et al. [11] has shown that in such contexts, first-person perspective avatars offer advantages over third-person perspective representations, particularly for local, asynchronous manual tasks. By aligning the avatar's movements directly with the user's own viewpoint, first-person guidance can facilitate accurate imitation, reduce ambiguity in joint positioning, and support efficient skill imitation. Further, the timing of first-person instructions not only affects the time needed but also the memorization of steps [12]. Therefore, correct imitation and timing may not only improve individual performance but also enhance safety for both users and nearby standers by improving memory, reducing procedural errors, and unintended movements in training scenarios.

However, in safety-critical and dynamic environments, a first-person perspective is not universally optimal. When users must follow precise paths between locations in unpredictable or hazardous environments, it can be crucial to first observe the intended path and movement strategy from an external viewpoint. In these cases, a third-person avatar can provide an overview of where and how to move, enabling users to anticipate obstacles, replicate trajectories more accurately, and plan their actions before execution [18, 12]. Such guidance may be derived from own or expert demonstrations [19], pre-programmed procedures, or AI-based analysis [20].

Therefore, avatars can function not only as instructional aids but also as predictive and preventive tools. By demonstrating ideal movements, visualizing safe trajectories, and highlighting potential hazards, avatars, whether representing humans or AI agents, can help users avoid dangerous actions. When combined with real-time environmental sensing and physiological monitoring, avatar behavior and representation can be dynamically adapted to the user's current state, supporting both attentiveness and emotional regulation in high-risk situations. Crucially, adapting the visualization and embodiment of avatars can elicit measurable differences in users' physiological and perceptual responses [8, 9]. Prior work has shown that avatar characteristics can influence pain perception [6, 7], behavior change [10], and even the risk of physical injury [21]. Therefore, adjusting factors such as body representation or visual prominence may improve environmental visibility and user comfort [1, 19, 12], but can simultaneously alter attention and risk perception.

Despite this, most existing studies evaluate guidance success at a broad level, focusing on whether and when task steps are completed correctly or efficiently. Little research has examined the fine-grained nuances of avatar-based guidance, like posture and movement trajectory, that are critical for preventing harm in dangerous situations. Understanding how these factors interact with avatar design and perspective remains an open challenge, particularly in environments where small errors can have severe consequences. Future work, therefore, needs to explore how to design instructions that are not only effective but also precise, safe, and context-aware.

3. Avatars as Social Mediators

Beyond individual instruction, avatars play a crucial role in collaborative XR scenarios, which include co-located users like teammates, remote collaborators, AI, and robot interaction, but also hybrid use cases. In such settings, avatars act as mediators of social presence, intention, and coordination, shaping how users perceive and work with others in shared spaces [22, 13].

Collaboration in CEs presents unique challenges. The social behavior of others, such as sudden movements, facial expressions, or perceived hesitation, can increase cognitive load, stress, and nervousness. In safety-critical situations, these social factors can directly impact performance and increase the likelihood of errors or accidents [23]. XR avatars offer a mechanism to modulate these social signals by adapting expressions, gestures, or movement timing to support smoother interaction and reduce interpersonal tension. Through controlled avatar representations, systems can decrease overly expressive behaviors, synchronize movements, or emphasize clear and predictable action cues. Such adaptations may help avoid conflict, improve mutual understanding, and maintain focus during high-risk collaborative tasks. While the modulation of social signals raises ethical considerations in terms of autonomy, deception, and the legitimacy of system-driven social engineering, it may prevent injury or loss of life. This underscores the need for careful design of avatar-mediated collaboration, where such measures should be jointly determined with users and remain adaptable to individual preferences.

Avatars also play a key role in human–robot collaboration. Overlaying a remote- or AI-controlled robot with an avatar representing the human operator [24], or another trusted individual [25], may increase users' sense of security and predictability. Such embodiment strategies can make robotic actions easier to interpret, foster trust, and reduce hesitation in time-critical tasks. While such visualizations can enhance safety, they may also introduce new risks. For instance, representing a robot as softer, lighter or less hazardous than it actually is can lead users to underestimate potential danger, increasing the likelihood of injury. For example, in the event of a malfunction, such as uncontrolled movement or a complete loss of responsiveness, the situation may be misinterpreted, which in the worst case can result in serious harm. To mitigate these risks, XR systems should incorporate adaptive safety mechanisms, such as dynamically adjusting visualizations and providing clear, behavior-dependent warning cues.

Taken together, avatar-based collaboration extends beyond representation to become an active design tool for regulating social interaction, trust, and coordination under stress. Understanding how avatar embodiment, behavior, and adaptation influence collaborative performance, harmony, and risk perception in CEs remains an important but still limited area for HCI research.

4. Conclusion

Advances in XR and AI enable avatars to evolve from static representations into adaptive agents that support guidance, collaboration, and safety in challenging environments. While prior work demonstrates clear benefits of avatar-based systems, our analysis highlights critical trade-offs between attention, trust, and situational awareness, particularly in CEs, where even minor errors can have severe consequences. We argue that avatar design must be inherently context-sensitive, dynamically balancing instructional support with unobstructed perception of the physical environment, while also regulating social signals to promote calm, attentive, and coordinated behavior. Addressing these challenges requires a deeper investigation into fine-grained aspects of movement, perspective, and embodiment, alongside careful consideration of ethical implications. Together, these insights outline a research agenda for the development of adaptive, human-centered avatar systems that enhance performance and safety without introducing new risks.

5. The Author

Clara Sayffaerth is a PhD student at LMU Munich with a background in Human–Computer Interaction and Mechanical Engineering. Her research explores the use of XR and AI-driven avatars for knowledge transfer. In collaboration with the Department of Physics Education and interdisciplinary partners, she develops and evaluates instructional XR systems that try to enhance learning and guidance while systematically identifying potential risks. Sayffaerth will continue her research [11, 12], particularly in the area of CEs, with the goal to create XR instructions that are efficient and safe while considering social aspects.

Declaration on Generative AI

During the preparation of this work, the author used ChatGPT and Grammarly in order to rephrase sentences, check grammar, and spelling. Further, the author used Adobe Firefly to generate parts of the teaser figure. After using these tools, the author reviewed and edited the content as needed and takes full responsibility for the publication's content.

References

[1] Y. Cao, X. Qian, T. Wang, R. Lee, K. Huo, K. Ramani, An Exploratory Study of Augmented Reality Presence for Tutoring Machine Tasks, in: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, ACM, Honolulu HI USA, 2020, pp. 1–13. URL: https://dl.acm.org/doi/10.1145/3313831.3376688. doi:10.1145/3313831.3376688.

[2] S. Thanyadit, P. Punpongsanon, T. Piumsomboon, T.-C. Pong, XR-LIVE: Enhancing Asynchronous Shared-Space Demonstrations with Spatial-temporal Assistive Toolsets for Effective Learning in Immersive Virtual Laboratories, Proceedings of the ACM on Human-Computer Interaction 6 (2022) 136:1–136:23. URL: https://dl.acm.org/doi/10.1145/3512983. doi:10.1145/3512983.

[3] H. Takita, K. Hashiura, Y. Hatada, D. Kodama, T. Narumi, T. Tanikawa, M. Hirose, Do We Still Need Human Instructors? Investigating Automated Methods for Motor Skill Learning in Virtual Co-Embodiment, IEEE Transactions on Visualization and Computer Graphics 31 (2025) 2455–2463. URL: https://ieeexplore.ieee.org/document/10918851/. doi:10.1109/TVCG.2025.3549540.

[4] J. Zhang, B. Han, Z. Dong, R. Wen, G. A. Lee, S. Hoermann, W. Zhang, T. Piumsomboon, Virtual Triplets: A Mixed Modal Synchronous and Asynchronous Collaboration with Human-Agent Interaction in Virtual Reality, in: Extended Abstracts of the CHI Conference on Human Factors in Computing Systems, CHI EA '24, Association for Computing Machinery, New York, NY, USA, 2024, pp. 1–8. URL: https://dl.acm.org/doi/10.1145/3613905.3650766. doi:10.1145/3613905.3650766.

[5] S. Aseeri, V. Interrante, The Influence of Avatar Representation on Interpersonal Communication in Virtual Social Environments, IEEE Transactions on Visualization and Computer Graphics 27 (2021) 2608–2617. URL: https://ieeexplore.ieee.org/abstract/document/9382845. doi:10.1109/TVCG.2021.3067783.

[6] M. Matamala-Gomez, A. M. Diaz Gonzalez, M. Slater, M. V. Sanchez-Vives, Decreasing Pain Ratings in Chronic Arm Pain Through Changing a Virtual Body: Different Strategies for Different Pain Types, The Journal of Pain 20 (2019) 685–697. URL: https://www.sciencedirect.com/science/article/pii/S1526590018309945. doi:10.1016/j.jpain.2018.12.001.

[7] M. Martini, D. Perez-Marcos, M. V. Sanchez-Vives, What color is my arm? Changes in skin color of an embodied virtual arm modulates pain threshold, Frontiers in human neuroscience 7 (2013) 438. doi:10.3389/fnhum.2013.00438.

[8] M. Kocur, F. Habler, V. Schwind, P. W. Woźniak, C. Wolff, N. Henze, Physiological and Perceptual Responses to Athletic Avatars while Cycling in Virtual Reality, in: Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems, ACM, Yokohama Japan, 2021, pp. 1–18. URL: https://dl.acm.org/doi/10.1145/3411764.3445160. doi:10.1145/3411764.3445160.

[9] M. Kocur, L. Jackermeier, V. Schwind, N. Henze, The Effects of Avatar and Environment on Thermal Perception and Skin Temperature in Virtual Reality, in: Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems, CHI '23, Association for Computing Machinery, New York, NY, USA, 2023, pp. 1–15. URL: https://dl.acm.org/doi/10.1145/3544548.3580668. doi:10.1145/3544548.3580668.

[10] K. Kilteni, I. Bergstrom, M. Slater, Drumming in Immersive Virtual Reality: The Body Shapes the Way We Play, IEEE Transactions on Visualization and Computer Graphics 19 (2013) 597–605. URL: https://ieeexplore.ieee.org/abstract/document/6479188. doi:10.1109/TVCG.2013.29.

[11] C. Sayffaerth, A. Köhler, J. Rasch, A. Schmidt, F. Müller, Through the Expert's Eyes: Exploring Asynchronous Expert Perspectives and Gaze Visualizations in XR, IEEE Computer Society, 2025, pp. 1311–1321. URL: https://www.computer.org/csdl/proceedings-article/ismar/2025/876100b311/2byAmwGMjW8. doi:10.1109/ISMAR67309.2025.00136.

[12] C. Sayffaerth, E. Ablimit, A. Köhler, J. Wombacher, A. Schmidt, F. Müller, Do It Fast, Forget It Fast: How Timing and Limb Visualizations Affect First-Person Augmented Reality Instructions, in: Proceedings of the 2026 CHI Conference on Human Factors in Computing Systems, CHI '26, Association for Computing Machinery, Barcelona, Spain, 2026, pp. 1–14. URL: https://dl.acm.org/doi/10.1145/3772318.3791471. doi:10.1145/3772318.3791471.

[13] D. I. Fink, M. Skowronski, J. Zagermann, A. V. Reinschluessel, H. Reiterer, T. Feuchtner, There Is More to Avatars Than Visuals: Investigating Combinations of Visual and Auditory User Representations for Remote Collaboration in Augmented Reality, Proceedings of the ACM on Human-Computer Interaction 8 (2024) 540–568. URL: https://dl.acm.org/doi/10.1145/3698148. doi:10.1145/3698148.

[14] J. Sutton, T. Langlotz, A. Plopski, K. Hornbæk, Flicker Augmentations: Rapid Brightness Modulation for Real-World Visual Guidance using Augmented Reality, in: Proceedings of the CHI Conference on Human Factors in Computing Systems, CHI '24, Association for Computing Machinery, New York, NY, USA, 2024, pp. 1–19. URL: https://dl.acm.org/doi/10.1145/3613904.3642085. doi:10.1145/3613904.3642085.

[15] A. Tang, C. Owen, F. Biocca, W. Mou, Comparative effectiveness of augmented reality in object assembly, in: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, ACM, Ft. Lauderdale Florida USA, 2003, pp. 73–80. URL: https://dl.acm.org/doi/10.1145/642611.642626. doi:10.1145/642611.642626.

[16] A. Khokhar, C. Borst, Modifying Pedagogical Agent Spatial Guidance Sequences to Respond to Eye-Tracked Student Gaze in VR, in: Proceedings of the 2022 ACM Symposium on Spatial User Interaction, ACM, Online CA USA, 2022, pp. 1–12. URL: https://dl.acm.org/doi/10.1145/3565970.3567697. doi:10.1145/3565970.3567697.

[17] T. Piumsomboon, G. A. Lee, A. Irlitti, B. Ens, B. H. Thomas, M. Billinghurst, On the Shoulder of the Giant: A Multi-Scale Mixed Reality Collaboration with 360 Video Sharing and Tangible

Interaction, in: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, CHI '19, Association for Computing Machinery, New York, NY, USA, 2019, pp. 1–17. URL: https://dl.acm.org/doi/10.1145/3290605.3300458. doi:10.1145/3290605.3300458.

[18] J. Friedman, M. Korman, Observation of an expert model induces a skilled movement coordination pattern in a single session of intermittent practice, Scientific Reports 9 (2019) 4609. URL: https://www.nature.com/articles/s41598-019-40924-9. doi:10.1038/s41598-019-40924-9.

[19] E. Lampen, J. Lehwald, T. Pfeiffer, A Context-Aware Assistance Framework for Implicit Interaction with an Augmented Human, in: J. Y. C. Chen, G. Fragomeni (Eds.), Virtual, Augmented and Mixed Reality. Industrial and Everyday Life Applications, volume 12191, Springer International Publishing, Cham, 2020, pp. 91–110. URL: https://link.springer.com/10.1007/978-3-030-49698-2_7. doi:10.1007/978-3-030-49698-2_7, series Title: Lecture Notes in Computer Science.

[20] N. Zheng, X. Song, T. Su, W. Liu, Y. Yan, L. Nie, Egocentric Early Action Prediction via Adversarial Knowledge Distillation, ACM Transactions on Multimedia Computing, Communications, and Applications 19 (2023) 1–21. URL: https://dl.acm.org/doi/10.1145/3544493. doi:10.1145/3544493.

[21] A. Cheymol, R. Fribourg, A. Lécuyer, J.-M. Normand, F. Argelaguet, Beyond my Real Body: Characterization, Impacts, Applications and Perspectives of "Dissimilar" Avatars in Virtual Reality, IEEE Transactions on Visualization and Computer Graphics 29 (2023) 4426–4437. URL: https://ieeexplore.ieee.org/document/10269053/. doi:10.1109/TVCG.2023.3320209.

[22] T. Piumsomboon, A. Dey, B. Ens, G. Lee, M. Billinghurst, The Effects of Sharing Awareness Cues in Collaborative Mixed Reality, Frontiers in Robotics and AI 6 (2019). URL: https://www.frontiersin.org/journals/robotics-and-ai/articles/10.3389/frobt.2019.00005/full. doi:10.3389/frobt.2019.00005.

[23] S. G. Barsade, The Ripple Effect: Emotional Contagion and its Influence on Group Behavior, Administrative Science Quarterly 47 (2002) 644–675. URL: http://journals.sagepub.com/doi/10.2307/3094912. doi:10.2307/3094912.

[24] B. Jones, Y. Zhang, P. N. Y. Wong, S. Rintel, Belonging There: VROOM-ing into the Uncanny Valley of XR Telepresence, Proc. ACM Hum.-Comput. Interact. 5 (2021) 59:1–59:31. URL: https://dl.acm.org/doi/10.1145/3449133. doi:10.1145/3449133.

[25] Z. Yu, G. Venture, Enhancing Robot Expressiveness with Augmented Reality Avatar, in: Proceedings of the 13th International Conference on Human-Agent Interaction, ACM, Yokohama Japan, 2025, pp. 470–472. URL: https://dl.acm.org/doi/10.1145/3765766.3765836. doi:10.1145/3765766.3765836.