Enhancing Asynchronous Learning in immersive Environments: Exploring Baseline Modalities for Avatar-Based AR Guidance

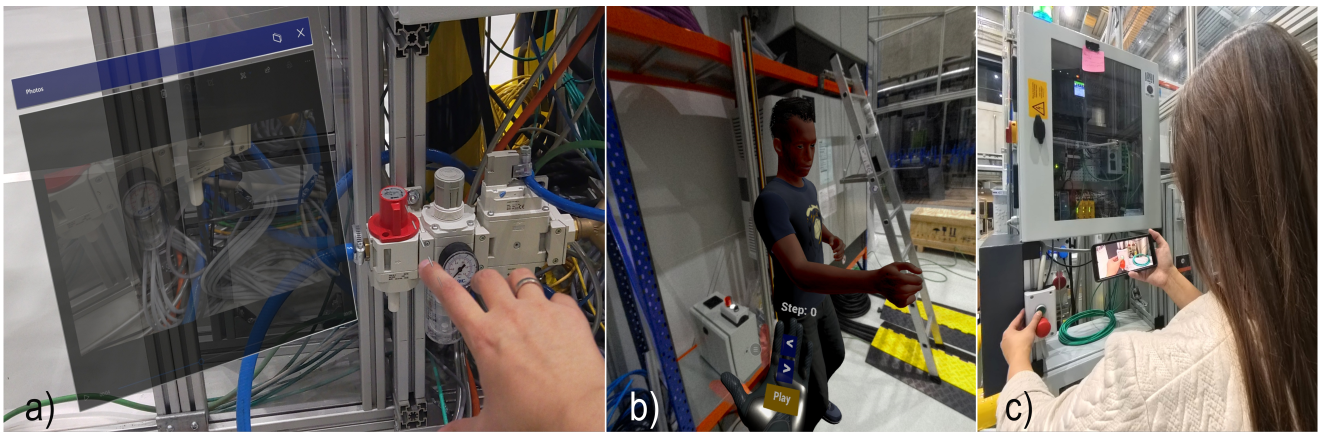
Anjela Mayer*
Karlsruhe Institute of Technology

Ines Miguel-Alonso†
University of Burgos

Jean-Rémy Chardonnet ‡
Arts et Métiers Institute of Technology, LISPEN

Andres Bustillo§
University of Burgos

Jivka Ovtcharova¶
Karlsruhe Institute of Technology



Tested guidance modalities in commissining task of a production cell: a) Augmented Reality video playback in a HoloLens, b) Augmented Reality avatar playback in a Meta Quest Pro device, and c) Handheld video playback on a smartphone.

Abstract

This study investigates baseline modalities for evaluating Augmented Reality (AR) avatar guidance in asynchronous collaboration on spatially complex tasks. A formative study with three participants compared smart-phone video, HoloLens video, and AR avatars across usability, collabora-tion, learning, and spatial awareness. Results suggest smartphone video as a reliable baseline due to usability and familiarity. Avatars showed potential for enhancing spatial awareness, task engagement, and learning outcomes but require interface improvements. Despite the small sample size, this study offers insights into immersive technologies for industrial training and collaboration.

Augmented Reality, Virtual Reality, Avatar, Asynchronous Collabo-ration, Learning, Training, Comparative Study.

*e-mail: anjela.mayer@kit.edu
†e-mail: imalonso@ubu.es
‡e-mail: jean-remy.chardonnet@ensam.eu
§e-mail: abustillo@ubu.es
¶e-mail: jivka.ovtcharova@kit.edu

1 Introduction and Motivation

Asynchronous collaboration, where users work on shared tasks at dif-ferent times, is gaining popularity in domains such as engineering and industrial training, where Extended Reality (XR) technologies have been recognized as transformative for improving teamwork and task coordi-nation [3]. In such domains, spatially complex tasks often require flex-ible, on-demand guidance that traditional methods (e.g. static instruc-tions, videos) struggle to provide. XR technologies, especially Augmented Reality (AR), can address the challenges of complex tasks by integrating immersive, context-rich environments into the real world, which im-proves collaboration and learning by leveraging 3D spatial information and natural user behaviours [4]. Unlike conventional approaches, AR sys-tems deliver dynamic, spatially aware instructions that can improve task comprehension, engagement, and knowledge retention. Despite its po-tential, asynchronous XR collaboration remains underexplored, with most research focusing on synchronous methods [8].

Even when asynchronous XR systems are developed, they are often validated in isolation without comparative benchmarking [1]. This lack of comparative studies limits our understanding of how novel approaches perform relative to more established methods like video instructions, particu-larly collaboration and learning outcomes. Research suggests that XR spa-tial instructions offer advantages over static methods by enhancing spatial interactions and stimulating collaborative presence, especially for com-plex tasks [6]. However, whether avatar-based guidance is superior to video-based guidance remains unclear. Previous work identifies video-based modalities as common approaches for task support, offering familiar and accessible formats on handheld devices (e.g., smartphones) or immer-sive playback through devices like HoloLens [2]. Research has also shown that avatar-based methods provide enhanced co-presence and engagement compared to video, and thus would improve learning experiences [9].

This initial study aims to find the most suitable video baseline for up-coming experiments on asynchronous collaboration in spatially complex, by assessing three guidance conditions:

- **AR Avatar**: Step-by-step task guidance by a virtual character using Meta Quest Pro, as an immersive condition.
- **AR Video**: Continuous first-person video on HoloLens, as a 2D video in an im-mersive space condition.
- **Handheld Video**: Continuous first-person video on a smartphone, providing a familiar and accessible condition.

2 Study Design and Procedure

Three participants (two males, one female) evaluated all three methods by performing six tasks in a robotic production environment as depicted in Figure . The video instructions were created using the built-in cam-era features of HoloLens and smartphones, while the avatar guidance was captured using an earlier developed system for avatar replay [5].

The tasks included switching on electrical cabinets, operating a robot control panel, adjusting the production piece, and verifying safety mea-sures (such as closing the door) before activating the robots. Three tasks (checking the laser barrier, adjusting the piece, and closing the door) in-tentionally included errors, requiring participants to record questions for asynchronous troubleshooting. This design enabled the evaluation of us-ability, engagement, and task comprehension across the three modalities.

The procedure began with a pretest to assess participants' baseline knowledge, followed by orientation on the technologies. To reduce nov-elty effect, AR familiarization was included [7]. Then, the guided tasks were carried out with each modality. Finally, participants completed ques-tionnaires to evaluate usability and user experience (SUS, UEQ), as well as collaboration, assessed through ten custom questions using a 5-point Likert scale. The custom metrics for collaboration measured coordination of actions, clarity of communication, and overall effectiveness of cooperation, but requires further validation to ensure reliability in diverse contexts. Qualitative feedback was also collected.

3 Results and Discussion

This study aimed to identify the most suitable video baseline for evaluation of asynchronous AR avatar collaboration in spatial tasks.

Usability and User Experience: A repeated-measures ANOVA on SUS scores revealed a significant difference between conditions ($F(2,4) = 20,8$, $p < 0.01$, $\eta_p^2 = 0.91$). Pairwise t-tests showed that the smartphone ($M = 90.8$, $SD = 13.8$) was significantly more usable than the HoloLens ($M = 55$, $SD = 9.0$, $t(2) = -4.51$, $p < 0.05$) , but not significantly more usable than the avatar ($M = 63.3$, $SD = 7.6$, $t(2) = -2,40$, $p = 0.14$). No significant difference was found between the avatar and HoloLens ($t(2) = 1,65$, $p = 0.24$). Participants attributed the smartphone's usability to its familiarity and simplicity, while the avatar's hands-free interaction was seen as beneficial despite technical issues. The UEQ indicated the highest stimulation and novelty scores for the avatar, though no significant differences were found.

Collaboration Metrics: Custom collaboration metrics indicated im-provements in communication, cooperation, and coordination in the avatar condition. A repeated-measures ANOVA revealed a significant difference in cooperation scores ($F(2,4) = 32,91$, $p < 0.01$, $\eta_p^2 = 0.94$). Pairwise
t-tests showed that HoloLens ($M = 2.9$, $SD = 0.84$) was significantly out-performedby both the smartphone ($M = 4.1$, $SD = 1.02$, $t(2) = -11.00$, $p < 0.01$) and the avatar ($M = 4.6$, $SD = 0.51$, $t(2) = 8.66$, $p < 0.05$), while no significant difference was found between the avatar and smart-phone ($t(2) = 1.51$, $p = 0.27$). Participants appreciated the avatar's spatial task demonstrations but noted that the smartphone's straightforward play-back also supported effective collaboration. Further validation of these metrics is needed.

Knowledge Retention: Pre- and post-test scores showed a 10% im-provement across all conditions. A one-tailed t-test indicated a significant increase in post-test scores ($t(2) = -3.46$, $p < 0.05$), suggesting improved knowledge retention. This suggests that while the avatar condition may enhance engagement and spatial presence, it may not significantly impact knowledge retention compared to traditional video formats.

Qualitative Feedback: Smartphone was described as *"simple and fa-miliar"*, but the avatar was appreciated for its spatial demonstration and immersive quality. Meta Quest Pro was favored over HoloLens due to visibility issues with holograms. Step-by-step guidance was preferred for asynchronous tasks, allowing participants to follow instructions at their own pace and revisit specific steps when needed. However, they also ex-pressed concerns about potential distraction and occlusion caused by the avatar, which could pose safety risks in environments where situational awareness is crucial. Participants felt a greater sence of immersion and spatial presence and reported that they could "step into" the avatar's per-spective to understand spatial relationships better, which could enhance task comprehension in complex setups.

4 Implications and Conclusion

These findings provide guidance for refining the design of the main study. The smartphone video condition scored high in usability and is consid-ered reliable for non-immersive baseline tasks due to its straightforward functionality. The AR condition, with its immersive and hands-free qual-ities, could enhance task comprehension and spatial awareness, but needs interface improvements. Conversely, the HoloLens condition had lower usability and user engagement scores, with participants citing visibility and interaction issues, suggesting it may be unsuitable for this context. The mentioned issues like bugs can be corrected in the future with the im-provement of HMDs tracking user body and lighting conditions, avoiding reflections and dim lighting environments. While the limited sample size restricted the analysis scope, this formative study highlights the potential for AR avatars in asynchronous collaboration. Future research should in-corporate a larger, more diverse sample, and further validate the custom collaboration metrics to strengthen the findings and support the develop-ment of effective XR-based learning tools.

This study was supported by the Ministry of Science, Innovation and Universities (FPU21/01978).

References

[1] A. Becker and C. M. D. S. Freitas. Evaluation of xr applications: a tertiary review. *ACM Comp. Surveys*, 2023. doi: 10.1145/3626517

[2] D. M. Calandra, A. Cannavò, and F. Lamberti. Improving ar-powered remote as-sistance: a new approach aimed to foster operator's autonomy and optimize the use of skilled resources. *The Int. Journal of Advanced Manufacturing Technology*, 114:3147 – 3164, 2021.

[3] B. Ens, J. Lanir, A. Tang, S. Bateman, G. Lee, T. Piumsomboon, and M. Billinghurst. Revisiting collaboration through mixed reality: The evolution of groupware. *International Journal of Human-Computer Studies*, 131, 2019. doi: 10.1016/j.ijhcs.2019.05.011

[4] A. R. Fender and C. Holz. Causality-preserving asynchronous reality. In *Proceed-ings of the 2022 CHI Conference on Human Factors in Computing Systems*, CHI '22. Association for Computing Machinery, New York, NY, USA, 2022. doi: 10. 1145/3491102.3501836

[5] K. Kastner, I. Miguel-Alonso, A. Mayer, J.-R. Chardonnet, A. Bustillo, J. Ovtcharova, R. Eiswirth, and J. Reichwald. Evaluating the effectiveness of avatar-based collaboration in xr for pump station training scenarios. In *2024 IEEE 3rd German Educ. Conf. (GECon)*, pp. 1–7, 2024. doi: 10.1109/GECon62014. 2024.10734032

[6] A. Mayer, A. Rungeard, J.-R. Chardonnet, P. Häfner, and J. Ovtcharova. Immersive hand instructions in ar – insights for asynchronous remote collaboration on spatio-temporal manual tasks. In *2023 IEEE Int. Conf. on Computational Intelligence and Virtual Environments for Measurement Systems and Applications (CIVEMSA)*, pp. 1–6, 2023. doi: 10.1109/CIVEMSA57781.2023.10231018

[7] I. Miguel-Alonso, D. Checa, H. Guillen-Sanz, and A. Bustillo. Evaluation of the novelty effect in immersive Virtual Reality learning experiences. *Virtual Reality*, 28(1):27, 2024. doi: 10.1007/s10055-023-00926-5

[8] C. Pidel and P. Ackermann. Collaboration in virtual and augmented reality: A systematic overview. In L. T. De Paolis and P. Bourdot, eds., *Augmented Real-ity, Virtual Reality, and Computer Graphics*, pp. 141–156. Springer International Publishing, Cham, 2020.

[9] S. Thanyadit, P. Punpongsanon, T. Piumsomboon, and T.-C. Pong. Xr-live: En-hancing asynchronous shared-space demonstrations with spatial-temporal assistive toolsets for effective learning in immersive virtual laboratories. *Proc. ACM Hum.-Comput. Interact.*, 6(CSCW1), Apr. 2022. doi: 10.1145/3512983