SpatialTutorX: Exploring Shared AI Facilitation for Collaborative Procedural Learning in Mixed Reality

Duo Wang
College of Medicine
University of Illinois
Urbana-Champaign
Champaign, Illinois, USA
dw36@illinois.edu

Jianwei Ni
College of Liberal Arts & Sciences
University of Illinois at
Urbana-Champaign
Champaign, Illinois, USA
ni12@illinois.edu

Qingxiao Zheng
College of Information Sciences and
Technology
The Pennsylvania State University
State College, Pennsylvania, USA
qpz5159@psu.edu

Shandra Jamison
College of Medicine
University of Illinois
Urbana-Champaign
Champaign, Illinois, USA
shandraj@illinois.edu

Mike Yao
Institute of Communications Research
University of Illinois
Urbana-Champaign
Champaign, Illinois, USA
Mzyao@illinois.edu

Abstract

Professional procedural training often requires learners to coordinate attention, communication, and action around shared physical artifacts, while expert facilitation remains limited, inconsistent, and difficult to scale. In real-world medical training and practice, clinicians frequently collaborate with peers, negotiate procedural decisions, and coordinate around shared physical tasks. We propose SpatialTutorX, a mixed-reality procedural training system that explores shared AI-MR facilitation embedded within immersive collaborative procedural work. Building on our prior single-user system for lumbar puncture training, this new version enables two learners wearing mixed-reality headsets to participate in the same training session around a shared physical manikin. Learners view synchronized mixed-reality procedural guidance, coordinate actions around the manikin, discuss procedural decisions with one another, and jointly interact with the same AI facilitator during the task. Through this demo, we invite CSCW attendees to examine how AI agents can function as shared instructional actors within collaborative, embodied, and socially organized professional learning practices.

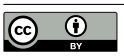
CCS Concepts

• **Human-centered computing → Mixed / augmented reality**.

Keywords

Mixed Reality, Collaborative Learning, Shared AI Facilitator, Human-AI Interaction, Medical Simulation

ACM Reference Format:
Duo Wang, Jianwei Ni, Qingxiao Zheng, Shandra Jamison, and Mike Yao. 2026. SpatialTutorX: Exploring Shared AI Facilitation for Collaborative Procedural Learning in Mixed Reality. In *Companion of the Computer-Supported Cooperative Work and Social Computing (CSCW Companion '26), October 10–14, 2026, Salt Lake City, UT, USA.* ACM, New York, NY, USA, 3 pages. https://doi.org/10.1145/3785651.3831422



This work is licensed under a Creative Commons Attribution 4.0 International License.
CSCW Companion '26, Salt Lake City, UT, USA
© 2026 Copyright held by the owner/author(s).
ACM ISBN 979-8-4007-2378-0/26/10
https://doi.org/10.1145/3785651.3831422

1 Introduction

Many forms of professional training are collaborative, embodied, and situated around shared physical work [3]. In medical education, for example, procedural learning often requires learners to coordinate attention, interpretation, communication, and physical actions around shared artifacts while relying on instructors, simulation facilitators, or peers for guidance. Simulation-based procedural training is often constrained by limited instructor availability, variability across instructors and sessions, and insufficient real-time feedback during practice [1, 9]. Learners may need help at the exact moment when their hands, attention, and reasoning are already occupied by the procedure. Mixed reality (MR) and AI offer new possibilities for supporting collaborative procedural work. MR can anchor instructional content directly onto physical training artifacts, while AI can provide on-demand explanations, procedural reminders, and conversational support [2, 6].

We propose to demonstrate SpatialTutorX, a mixed-reality procedural training system that explores shared AI-MR facilitation in immersive collaborative procedural work through the use case of lumbar puncture training. The original version of *SpatialTutor* explored how mixed reality and AI could support procedural learning through object-aware spatial guidance, spatially anchored demonstrations, and conversational AI assistance [11]. This CSCW adaptation builds on prior work on proactive versus reactive AI facilitation in MR medical training, which examined how the timing and initiative of AI support can shape learner experience in MR procedural learning [12]. In this demo, we extend *SpatialTutor* into SpatialTutorX, a multiplayer collaborative learning environment in which two medical trainees wearing Apple Vision Pro headsets participate in the same procedural training session around a shared physical manikin, synchronized mixed-reality guidance, and a common AI facilitator. Rather than positioning AI as an individual tutor, SpatialTutorX embeds AI facilitation directly into peer collaborative

activity, where learners jointly coordinate procedural actions, negotiate decisions, and interact with the same AI facilitator throughout the task.

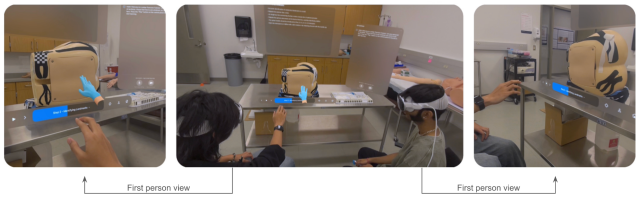


Figure 1: Illustration of the SpatialTutorX. Two learners wearing Apple Vision Pro headsets collaboratively engage with the same physical lumbar puncture manikin while viewing synchronized mixed-reality guidance.

2 System Overview

SpatialTutorX is organized around a shared procedural workspace that combines a physical lumbar puncture manikin, synchronized MR guidance, and shared AI facilitation. Lumbar puncture serves as a representative procedural task because it requires anatomical localization, careful hand positioning, sequential execution, and collaborative decision-making under uncertainty [5, 7].

The system supports a shared procedural state across two learners wearing Apple Vision Pro headsets, see Figure 1. First, learners view synchronized mixed-reality overlays anchored directly onto the physical manikin and procedural tray, where anatomical cues, instrument labels, and step-specific visual prompts appear near relevant physical objects. This allows both learners to remain aware of the same procedural state while connecting abstract procedural knowledge with the material workspace. Second, spatially anchored 3D demonstrations show procedural movements, tool manipulation, and anatomical orientation directly over the physical task trainer, providing situated procedural references that learners can simultaneously observe, revisit, and discuss together during practice. Third, the system includes a shared conversational AI facilitator that provides step-aware explanations and procedural assistance through voice or text interaction. The AI can return multimodal responses, including spoken guidance, written summaries, and relevant instructional cues, which become shared conversational resources that both learners can jointly interpret, negotiate, and act upon during the procedure.

In this demonstration, two learners participate in the same procedural training session around a shared physical manikin while viewing synchronized instructional content and remaining aware of the same procedural state throughout the task. This design foregrounds the cooperative activity of learning around a shared artifact: learners must coordinate attention between the physical manikin, MR overlays, each other, and the AI facilitator while jointly discussing procedural decisions and interpreting AI guidance together. In this demo version, we focus on making the shared training session reliable, repeatable, and hands-on for conference attendees.



Figure 2: Synchronized MR guidance in SpatialTutorX



Figure 3: User interacting with manikin in MR

3 Cooperative Activities Supported by the Demo

The demo supports cooperative procedural learning around a shared manikin, as shown in Figure 2 and Figure 3. Two attendees will jointly engage with a simplified lumbar puncture learning sequence. They will inspect the manikin and instrument tray, interpret synchronized MR guidance, discuss steps, and decide when to ask the shared AI facilitator for help. The purpose is not to train attendees to perform lumbar puncture, but to let them experience a concrete scenario in which AI, MR, physical artifacts, and peer interaction are intertwined in collaborative professional learning.

The planned demo interaction follows a short, repeatable loop suitable for the conference demo setting. First, presenters introduce the training context and help two attendees enter the MR session. Second, attendees view the shared manikin-based training scene and receive a brief orientation to the synchronized MR overlays. Third, the pair completes one or two representative collaborative micro-tasks, such as identifying anatomical landmarks, discussing the next procedural step, or interpreting a spatial demonstration. Fourth, attendees ask the AI facilitator a question, such as where to locate the insertion point, what a specific instrument is used for, or what to do if they are uncertain about the next step. Finally, attendees are invited to reflect on how the shared AI facilitator shaped the collaborative activity: who initiated interaction with the AI, how the pair negotiated questions, and whether the AI felt like a tutor, teammate, facilitator, evaluator, or interface element.

4 Contributions to the CSCW Community

SpatialTutorX demonstrates a concrete implementation of AI-mediated cooperative activity in an immersive environment, connecting long-standing concerns around shared context, coordination, awareness, and the organization of cooperative work with emerging questions about AI participation in collaborative practice [3, 4, 8, 10]. Rather than positioning AI as a background model or private tutoring interface, SpatialTutorX embeds a shared AI facilitator directly within collaborative embodied procedural work. This allows the demo to foreground how AI participation reshapes peer interaction, shared decision-making, and instructional coordination around physical artifacts.

First, it provides a working prototype for discussing how AI agents should participate in collaborative learning. In SpatialTutorX, the AI is not merely a background model that generates content; it is an audible and conversational participant in the learning activity. This creates opportunities to discuss AI roles, such as tutor, coach, teammate, facilitator, evaluator, or social actor.

Second, the demo makes visible the coordination challenges that emerge when immersive guidance, physical artifacts, peer learning,

and AI facilitation converge. Learners must decide where to look, whose guidance to follow, when to speak, when to act, and how to interpret AI responses in relation to peer discussion. These are classic issues in shared context but situated in an immersive training environment.

Third, SpatialTutorX offers a bridge between research communities. The system builds from work in MR, medical simulation, and AI-supported education, but the demo foregrounds the social organization of learning and the collaborative role of AI. We believe this demo will generate productive conversations with CSCW attendees about how to design AI agents that support, rather than replace or disrupt, collaborative professional practice, especially when learning and work are embodied, high-stakes, and socially organized.

References

[1] Jeffrey H Barsuk, Elaine R Cohen, Joe Feinglass, William C McGaghie, and Diane B Wayne. 2009. Use of simulation-based education to reduce catheter-related bloodstream infections. *Archives of internal medicine* 169, 15 (2009), 1420–1423.

[2] Subramanian Chidambaram, Hank Huang, Fengming He, Xun Qian, Ana M Villanueva, Thomas S Redick, Wolfgang Stuerzlinger, and Karthik Ramani. 2021. Processar: An augmented reality-based tool to create in-situ procedural 2d/3d ar instructions. In *Proceedings of the 2021 ACM Designing Interactive Systems Conference*. 234–249.

[3] Paul Dourish and Victoria Bellotti. 1992. Awareness and coordination in shared workspaces. In *Proceedings of the 1992 ACM conference on Computer-supported cooperative work*. 107–114.

[4] Jonathan Grudin. 1994. Computer-supported cooperative work: History and focus. *Computer* 27, 5 (1994), 19–26.

[5] Anguraj Jagadish, Srinivasan Swaminathan, Prasanna U Bidkar, Suman L Gupta, and Sethuramachandran Adinarayanan. 2021. Ease of lumbar epidural catheter insertion with prepuncture ultrasound as guidance compared with conventional palpatory technique when performed by anesthesiology residents: a randomized controlled trial. *Journal of Anaesthesiology Clinical Pharmacology* 37, 2 (2021), 216–220.

[6] Aaron Kotranza, D Scott Lind, and Benjamin Lok. 2011. Real-time evaluation and visualization of learner performance in a mixed-reality environment for clinical breast examination. *IEEE transactions on visualization and computer graphics* 18, 7 (2011), 1101–1114.

[7] Chun Li, Miaomiao Li, Yixian Wang, Shaoyi Li, Lin Cong, and Weining Ma. 2023. A controlled lumbar puncture procedure improves the safety of lumbar puncture. *Frontiers in Neuroscience* 17 (2023), 1304150.

[8] Jeremy Roschelle and Stephanie D Teasley. 1995. The construction of shared knowledge in collaborative problem solving. In *Computer supported collaborative learning*. Springer, 69–97.

[9] Taylor Sawyer, Marjorie White, Pavan Zaveri, Todd Chang, Anne Ades, Heather French, JoDee Anderson, Marc Auerbach, Lindsay Johnston, and David Kessler. 2015. Learn, see, practice, prove, do, maintain: an evidence-based pedagogical framework for procedural skill training in medicine. *Academic Medicine* 90, 8 (2015), 1025–1033.

[10] Kjeld Schmidt and Liam Bannon. 1992. Taking CSCW seriously: Supporting articulation work. *Computer supported cooperative work (CSCW)* 1, 1 (1992), 7–40.

[11] Duo Wang, Jianwei Ni, Yuhan Zhou, Weiyu Ding, Wenxuan Song, Shandra Jamison, Mike Yao, and Caroline Cao. 2026. SPATIALTUTOR: Object-Aware Mixed Reality Training for Procedural Medical Skills Training with AI-Driven Support. In *2026 IEEE International Conference on Artificial Intelligence and eXtended and Virtual Reality (AIxVR)*. IEEE, 57–66.

[12] Duo Wang, Wenxuan Song, Jianwei Ni, Qingxiao Zheng, Yuhan Zhou, Kyrian Liang, Mike Yao, and Caroline GL Cao. 2026. Should the AI Speak First? Evaluating Proactive vs. Reactive Facilitation in Mixed-Reality Medical Training. In *Proceedings of the 2026 CHI Conference on Human Factors in Computing Systems*. 1–19.