

Assessing Cognitive Extension in a Conversational AI-Powered Extended Reality Learning Environment

Tudor Bumbu

Ion Creangă State Pedagogical
University of Chişinău
Chişinău, Republic of Moldova
tudor.bumbu@math.md

Mihail Cuznetov

Faculty of Computers,
Informatics and
Microelectronics Technical
University of Moldova (TUM)
Chişinău, Republic of Moldova
mihail.cuznetov@utm.md

Olesea Caftanatov

V. A. Institute of Mathematics
and Computer Science MSU
Chişinău, Republic of Moldova
olesea.caftanatov@math.md

Tatiana Veverița

Ion Creangă State Pedagogical
University of Chişinău
Chişinău, Republic of Moldova
veverita.tatiana@upsc.md

Laura-Bianca Bilius

MintViz Lab, Stefan cel Mare
University of Suceava
Suceava, Romania
laura.bilius@usm.ro

ABSTRACT

Artificial intelligence (AI) and extended reality (XR) are increasingly being integrated to support the development of adaptive learning environments. However, evaluations of such systems mainly focus on usability and workload rather than on the extent to which they support users' cognitive processes. In this paper, we evaluate a conversational, AI-powered XR learning environment in which learners use a mixed-reality headset to view and inspect life-size 3D animal models embedded in their physical surroundings and ask spoken questions about them, with responses generated by a large language model (LLM). Seven adult participants completed the XR Cognitive Extension Scale (XRCES) and System Usability Scale (SUS) after interacting with the system. The indicators of perceived cognitive extension revealed an overall XRCES score of $M=3.8$, with the highest scores for perceptual and embodiment extension and comparatively lower scores for attention extension and cognitive integration. Based on the results reported at the dimension level, we derive preliminary XRCES-informed design implications for conversational XR learning environments.

AUTHOR KEYWORDS

Cognitive extension; extended reality; mixed reality; augmented reality; large language models; conversational interfaces; adaptive learning; XRCES.

DOI: 10.37789/icusi.2026.27

INTRODUCTION

Extended reality (XR) technologies are increasingly adopted in education due to immersive experiences that can support user engagement and learning [10,17]. Recent advances in large language models (LLMs) have also made natural language interaction a practical approach for learning applications [13], that enable users to ask questions and receive contextually relevant answers. Therefore, the integration of XR and conversational AI can support adaptive learning environments, while conversational AI serves as the generative component within an AI-in-the-loop interaction pattern [9,18].

However, evaluations of XR learning applications are mostly focused on usability and workload using instruments such as the System Usability Scale (SUS) [5] and NASA-TLX [11]. These instruments provide limited insight into whether and to what extent these systems extend user's cognitive processes in the sense of the extended mind thesis [7].

To address this gap, the XR Cognitive Extension Scale (XRCES) [3] was recently introduced as a dedicated instrument with 18 items covering seven dimensions of cognitive extension. XRCES provides a means of examining the resulting cognitive-extension profile at the level of its individual dimensions. In this context, this paper investigates an AI-powered XR learning environment delivered through a mixed reality headset. As it remains unclear how users perceive cognitive extension in such environment, we use XRCES to characterize perceived cognitive extension across its seven dimensions.

The contributions of this paper are as follows:

- (1) A conversational, AI-powered XR learning environment for exploring life-size 3D animal models on a standalone mixed reality headset, integrating speech-to-text, LLM-based response generation, and text-to-speech interaction.
- (2) A preliminary user study applying XRCES and SUS to assess cognitive extension and usability in a conversational XR learning context, representing, to the best of our knowledge, the first assessment of cognitive extension with XRCES in a headset-based conversational setting.
- (3) A dimension-level analysis of the system's cognitive extension profile, from which we derive preliminary XRCES-informed design implications for conversational XR learning environments.

RELATED WORK

XR and AI in Education

XR technologies have been increasingly explored as a means of supporting immersive learning experiences, with prior work reporting benefits for learner engagement, motivation,

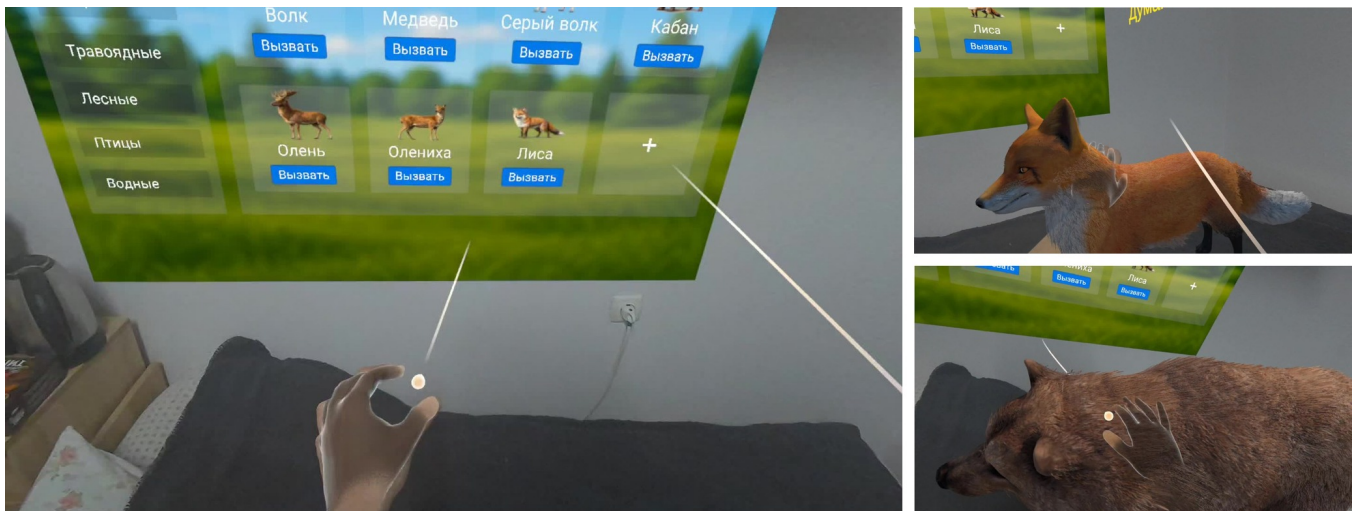


Figure 1: The prototype in use: animal cards grouped by categories and selected with hand-based interaction (left), a life-size fox placed in the physical environment (top right) and interaction with a life-size bear placed in the physical environment (bottom right).

and spatial understanding [10]. At the same time, their adoption remains subject to challenges related to content preparation, device limitations, and integration into educational settings [12]. Educational AR systems span a broad range of approaches, from marker-based applications tailored to different learning styles [12] and mobile AR applications for STEM education [4] to AI-based educational platforms incorporating speech-based interaction for primary education [23].

Recent work has begun to explore the integration of immersive technologies and generative AI. For example, Cheng *et al.* [6] integrated AR, narrative-based learning, and LLM-generated interactions in an educational application for elementary-school learners, illustrating the potential of LLMs to personalize immersive learning experiences. Vision-ARy [14] integrated ChatGPT with AR glasses to support contextualized language learning, showing how conversational AI can connect information delivery with the learner's physical surroundings. More recent systems have investigated embodied LLM agents in immersive environments. ConversAR [1] employed embodied LLM-powered agents in AR to support contextualized group conversations for second-language learning, while ELLMA-T [16] employed an embodied LLM tutor in social VR to provide situated conversational practice and adaptive feedback.

Measuring Cognitive Extension in XR

Several cognitive theories provide perspectives on how users process and interact with virtual content, including cognitive load theory, the 4E cognition framework, and the extended mind thesis [7]. Complementary evaluation instruments capture related dimensions of the XR experience, including cognitive load [11], presence [22], and perceived usability [5,15]. Related research has examined computational and critical thinking in XR-based learning [21], characterized the technological dimensions of XR-based cognitive augmentation systems [19], and proposed extended thinking as a conceptual framework for effective interaction in XR environments [20]. Prior work has also begun to consider how cognitive extension can be assessed in XR environments. In this context, XRCES [3] operationalizes cognitive extension as a measurable construct across seven dimensions, such as perceptual,

memory, attention, embodiment, thinking, cognitive integration, and cognitive sustainability. The instrument was validated in a between-subjects study ($N = 48$) using a mobile AR application for STEM learning, where it was shown to be sensitive to differences in the duration of XR immersion [3]. Our study applies XRCES in a different setting, *i.e.* a conversational, AI-powered XR learning environment implemented on a mixed reality headset.

THE CONVERSATIONAL XR LEARNING ENVIRONMENT

We introduce a prototype, an AI-powered XR learning environment designed to support learning about animals through spatial visualizations and natural-language interaction.¹ The application targets young learners (*i.e.*, children) and runs on a Meta Quest 3S² mixed reality headset with video see-through pass-through. Learners interact with a floating interface containing animal cards grouped by category. Selecting a card places a life-size, animated 3D model of the corresponding animal within the learner's physical surrounding; see Figure 1. The 3D model can be observed from different viewpoints as the learners move around it. Interaction is performed using hand gestures and voice, without handheld controllers, in order to simplify access to the system for childrens [9].

The system comprises three main components, such as a client application developed in Unity using the Meta XR SDK (pass-through visualization, spatial tracking, hand tracking), an ASP.NET Web API backend and an interchangeable LLM accessed through an HTTP API [9]. As can be seen in Figure 2, voice interaction follows a multi-stage processing pipeline. The learner's spoken question is first transcribed and sent from the client to the backend. The backend then incorporates the question into a prompt that constrains the generated response to remain concise, age-appropriate, and relevant to the currently selected animal. The final response is returned to the client and is presented to the user both visually and through text-to-speech [9].

¹The prototype was developed as part of a master's thesis project [9].

²<https://www.meta.com/quest/quest-3s/>

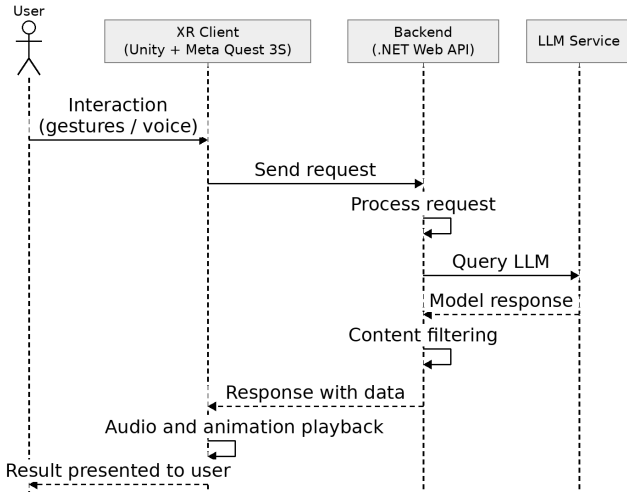


Figure 2: Sequence diagram illustrating voice interaction between the client (Unity on Meta Quest 3S), the ASP.NET backend responsible for prompt adaptation and content filtering, and the interchangeable LLM.

METHOD

We conducted a preliminary user study to characterize perceived cognitive extension across the seven XRCES dimensions in a conversational, AI-powered XR learning environment. In addition to XRCES, we collected usability ratings to complement the cognitive extension assessment.

Participants

A total of seven adults, four male and three female, aged between 20 and 40 years, were recruited for the preliminary user study. Participants were university students and staff members with varying levels of prior experience with XR technologies. Participation was voluntary, and informed consent was obtained from all participants before the study session.

Task

Each participant completed a ten minute exploration session, during which they explored at least three 3D animal models, inspected them from multiple angles, and asked at least five questions overall, with no other constraints. The session duration was selected for practical reasons, given the exploratory nature of this initial stage.

Measures and procedure

After the exploration session, participants completed the XRCES [3] (18 items rated on a 5-point scale from 1 = “very little” to 5 = “very strong”; see Table 1), followed by the SUS [5] questionnaire. Sessions were conducted individually in the same room and lasted approximately 20 minutes in total. For XRCES, dimension scores were calculated as the mean of the corresponding items, with the attention items reverse-scored. The overall XRCES score was calculated as the mean of the seven dimension scores, and internal consistency was assessed using Cronbach’s alpha [8].

PRELIMINARY RESULTS

Table 2 summarizes the preliminary results of our participants ratings. The overall XRCES score was $M = 3.8$, with an internal consistency of $\alpha = 0.86$, across the 18 items. The System

Table 1: The 18 XRCES items [3] grouped by dimension (shortened English wording of the Romanian version administered in the study).

Dimension and items

D1 Perceptual extension. (1) The virtual object was perceived as if it were real; (2) it occupied a concrete place in physical space; (3) it became more prominent to the senses than the physical surroundings.

D2 Memory extension. (4) The object is remembered as a 3D element that was explored rather than as a 2D image; (5) its information “stuck” naturally, without a deliberate effort to memorize.

D3 Attention extension (reverse-scored). (6) Effort to stay focused on the virtual experience while ignoring the physical surroundings; (7) effort to follow and understand the information delivered through the object; (8) effort to return attention to the object after interruptions.

D4 Embodiment extension. (9) The object responded directly to hand movements; (10) its reactions were fluid and predictable rather than artificial; (11) interacting with it felt like interacting with a real object.

D5 Thinking extension. (12) The task could be solved from visual cues in the virtual environment, without prior mental calculation or deduction; (13) the solutions felt like one’s own ideas even when guided by the application; (14) decisions were faster than with traditional (e.g., paper-based) methods.

D6 Cognitive integration. (15) The application clearly showed the steps for solving the task; (16) it helped detect and correct mistakes in real time.

D7 Cognitive sustainability. (17) The application made the task feel much easier than without it; (18) new information was learned through the virtual object.

Table 2: Average XRCES scores for each dimension.

Dimension	Mean (M)	Standard Deviation (SD)
D ₁ Perceptual extension	4.1	0.5
D ₂ Memory extension	3.9	0.6
D ₃ Attention extension	3.5	0.7
D ₄ Embodiment extension	4.0	0.6
D ₅ Thinking extension	3.8	0.5
D ₆ Cognitive integration	3.6	0.8
D ₇ Cognitive sustainability	3.7	0.6

Usability score (SUS) was $M = 78$ ($SD = 10$), indicating good perceived usability [15].

At the dimension level, the highest mean scores were observed for perceptual extension (D_1 , $M = 4.1$) and embodiment extension (D_4 , $M = 4.0$), which may be associated with the life-size pass-through presentation and hand-based interaction. Memory extension (D_2 , $M = 3.9$) and thinking extension (D_5 , $M = 3.8$) also received relatively high scores. Attention extension (D_3 , $M = 3.5$) had the lowest mean score, while cognitive integration (D_6 , $M = 3.6$) showed the greatest variability ($SD = 0.8$). Cognitive sustainability (D_7) yielded a mean score of $M = 3.7$, indicating that participants generally perceived the system as supportive of the learning task.

DISCUSSION AND FUTURE WORK

Our preliminary findings suggest that XRCES can be applied beyond the mobile AR context in which it was introduced [3] and evaluated [2], and provides an interpretable dimension-level profile for a headset-based conversational system. The dimension-level results point to several XRCES-informed design implications, which are discussed below in relation to

future iterations of conversational XR learning environments. Given the exploratory nature of the study and the small sample size, the design implications derived from the XRCES dimension scores should be considered preliminary.

Preliminary XRCES-informed design guidelines

Perceptual extension. According to our results, perceptual extension suggests that life-size spatial presentation and pass-through visualization are relevant design elements. Future designs could improve this dimension by maintaining stable spatial placement and appropriate scale of virtual objects, and ensuring that virtual animals remain visually prominent within the physical environment without unnecessarily obscuring the learner's surroundings.

Memory extension examined during our study suggests that exploring the animals as spatial 3D objects may have supported how participants retained information. Future iterations could build on this by incorporating opportunities for learners to recall information after exploring an animal model. For example, before presenting the information, the system could ask learners to remember a previously discussed characteristic, identify an animal based on previous information, or briefly explain what they learned during exploration.

Attention extension received the lowest mean score, therefore, maintaining focus within the mixed-reality experience appears to be an area requiring further design attention. This could include reducing unnecessary visual elements, keeping conversational exchanges concise, and providing subtle cues that help learners return their attention to the relevant animal or information after an interruption.

Embodiment extension, based on our study suggests that hand-based interaction and the spatial behavior of the virtual animals should remain central to the interaction design. However, future iterations could support higher embodiment by ensuring that hand-based actions produce immediate responses and by introducing simple forms of direct manipulation or context-appropriate reactions from the virtual animals.

Thinking extension could be improved by designing conversational interactions that encourage learners to reason before receiving an answer from the AI. For example, the system could invite learners to make a prediction about an animal, explain an observed feature, or formulate a possible answer before the LLM provides additional information. Follow-up questions could then build on the learner's response rather than simply supplying facts.

Cognitive integration received the second-lowest mean score, which suggest that future iterations could be improved by providing guidance and feedback to help learners identify and correct misunderstandings during the task. Moreover, participants reported that delays in the conversational pipeline disrupted the continuity of interaction, but these delays could be reduced through streaming responses or immediate feedback acknowledging the learner's input.

Cognitive sustainability mean score results suggest that future designs should preserve opportunities for active learner engagement. For example, the system could provide concise and context-relevant explanations when assistance is needed and avoid overly complete responses that remove the need for learners to explore and reflect. In this way, the conversational AI will support the acquisition of new information through interaction with the virtual content.

Limitations

The present study has several limitations, which also point to directions for future research. First, the current findings are based on a small sample of seven adult participants and should therefore be considered preliminary. Moreover, although the prototype is intended for children, the present study involved only adults, therefore, the findings cannot be generalized to the target population. Future studies will involve children in educational settings.

Second, the 10-minute interaction period may have been too short for participants to fully explore the system and may have influenced the observed cognitive-extension scores. Longer sessions are planned to examine how extended exposure affects the different XRCES dimensions.

Finally, the current study involved a single application, a single session, and no control condition, which prevents us from separating the contribution of the conversational AI component from that of the XR presentation. Future experiments should therefore include comparison conditions that allow these contributions to be examined independently.

CONCLUSIONS

This paper examined perceived cognitive extension in a conversational, AI-powered XR learning environment that integrates life-size 3D animal models, hand-based gesture interaction, and spoken dialogue with an LLM. Using XRCES together with SUS, we obtained a preliminary dimension-level profile of the experience, with higher scores for perceptual and embodiment extension and comparatively lower scores for attention extension and cognitive integration. These preliminary results show that XRCES can provide design-relevant insights beyond the mobile AR context in which it was originally introduced. Based on the collected data, we also derived preliminary XRCES-informed design implications for conversational XR learning environments.

ACKNOWLEDGMENTS

This work was supported by a grant of the Ministry of Education and Research, CCCDI-UEFISCDI, project number PN-IV-PCB-RO-MD-2024-0130, within PNCI IV. Tudor Bumbu and Tatiana Veverița also acknowledge support from the research and innovation project "New extended reality technology for cognitive augmentation based on the superimposition of interactive experiences," strategic direction "Innovative Technologies, Sustainable Energy, Digitalization," registered in the State Register of Projects in the field of Science and Innovation of the Republic of Moldova, project code 25.80013.5007.60ROMD. Olesia Caftanatov acknowledges support from the project Artificial Intelligence-based Educational Platform "Gugută," no. 25.80012.0807.21TC, Young Researchers programme (2025–2026), Moldova State University, V. Andrunachievici Institute of Mathematics and Computer Science.

References

- [1] Jad Bendarkawi, Ashley Ponce, Sean Mata, Aminah Aliu, Yuhua Liu, Lei Zhang, Amna Liaqat, Varun Rao, and Andrés Monroy-Hernández. 2025. ConversAR: Exploring Embodied LLM-Powered Group Conversations in Augmented Reality for Second Language Learners. *arXiv* (05 2025). <https://doi.org/10.48550/arXiv.2505.24000>

- [2] Andrei Braicov, Tatiana Veverița, Ilona Popovici, Mihail Terenti, and Radu-Daniel Vatavu. 2026. Assessing Cognitive Extension in Extended Reality: Initial Evidence from a Confirmatory Factor Analysis. In *Proceedings of the ACM Symposium on Spatial User Interaction (SUI '26)*. ACM, New York, NY, USA, 3 pages. <https://doi.org/10.1145/3822518.3839317>
- [3] Andrei Braicov, Tatiana Veverița, Ilona Popovici, Mihail Terenti, and Radu-Daniel Vatavu. 2026. XRCES: A multidimensional instrument for assessing cognitive extension in extended reality. In *Proceedings of the 2026 IEEE International Conference on Metrology for Extended Reality, Artificial Intelligence and Neural Engineering (MetroXRINE '26)*. 6 pages.
- [4] Andrei Braicov, Tatiana Veverița, Mihail Terenti, and Ivan Budanaev. 2026. Evaluating User Experience and Cognitive Load in AR-supported STEM Education. In *2026 International Conference on Development and Application Systems (DAS)*. 304–309. <https://doi.org/10.1109/DAS69882.2026.11553320>
- [5] John Brooke. 1996. SUS: A 'quick and dirty' usability scale. Chapter 21. In *Usability and Evaluation in Industry*. Vol. 189. 6 pages. <https://doi.org/10.1201/9781498710411>
- [6] Alan Y. Cheng, Meng Guo, Melissa Ran, Arpit Ranasaria, Arjun Sharma, Anthony Xie, Khuyen N. Le, Bala Vinaithirthan, Shihe (Tracy) Luan, David Thomas Henry Wright, Andrea Cuadra, Roy Pea, and James A. Landay. 2024. Scientific and Fantastical: Creating Immersive, Culturally Relevant Learning Experiences with Augmented Reality and Large Language Models. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems (CHI '24)*. Association for Computing Machinery, New York, NY, USA, Article 275, 23 pages. <https://doi.org/10.1145/3613904.3642041>
- [7] Andy Clark and David Chalmers. 1998. The Extended Mind. *Analysis* 58, 1 (1998), 7–19. <https://doi.org/10.1093/analys/58.1.7>
- [8] Lee J. Cronbach. 1951. Coefficient alpha and the internal structure of tests. *Psychometrika* 16, 3 (1951), 297–334. <https://doi.org/10.1007/BF02310555>
- [9] Mihail Cuznețov. 2026. *Medii de învățare adaptative bazate pe inteligența artificială și realitatea augmentată*. Master's thesis. Technical University of Moldova, Chișinău.
- [10] Juan Garzón. 2021. An Overview of Twenty-Five Years of Augmented Reality in Education. *Multimodal Technologies and Interaction* 5, 7 (2021). <https://doi.org/10.3390/mti5070037>
- [11] Sandra G. Hart and Lowell E. Staveland. 1988. Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In *Human Mental Workload*, Peter A. Hancock and Najmedin Meshkati (Eds.). Advances in Psychology, Vol. 52. North-Holland, 139–183. [https://doi.org/10.1016/S0166-4115\(08\)62386-9](https://doi.org/10.1016/S0166-4115(08)62386-9)
- [12] Titchiev Inga, Olesea Caftanator, and Dan Talambuta. 2023. Challenges associated with using AR technology in education. *Computer Science Journal of Moldova* 31 (12 2023), 367–380. <https://doi.org/10.56415/csjm.v31.19>
- [13] Enkelejda Kasneci, Kathrin Sessler, Stefan Küchemann, Maria Bannert, Daryna Dementieva, Frank Fischer, Urs Gasser, Georg Groh, Stephan Günnemann, Eyke Hüllermeier, Stephan Krusche, Gitta Kutyniok, Tilman Michaeli, Claudia Nerdel, Jürgen Pfeffer, Oleksandra Poquet, Michael Sailer, Albrecht Schmidt, Tina Seidel, Matthias Stadler, Jochen Weller, Jochen Kuhn, and Gjergji Kasneci. 2023. ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences* 103 (2023), 102274. <https://doi.org/10.1016/j.lindif.2023.102274>
- [14] Hyungmin Lee, Chen-Chun Hsia, Aleksandr Tsoy, Sungmin Choi, Hanchao Hou, and Shiguang Ni. 2023. VisionARy: Exploratory research on Contextual Language Learning using AR glasses with ChatGPT. In *Proceedings of the 15th Biannual Conference of the Italian SIGCHI Chapter (CHIItaly '23)*. Association for Computing Machinery, New York, NY, USA, Article 22, 6 pages. <https://doi.org/10.1145/3605390.3605400>
- [15] James R. Lewis and Jeff Sauro. 2018. Item benchmarks for the system usability scale. *Journal of Usability Studies* 13, 3 (May 2018), 158–167. <https://dl.acm.org/doi/abs/10.5555/3294033.3294037>
- [16] Mengxu Pan, Alexandra Kitson, Hongyu Wan, and Mirjana Prpa. 2025. ELLMA-T: an Embodied LLM-agent for Supporting English Language Learning in Social VR. In *Proceedings of the 2025 ACM Designing Interactive Systems Conference (DIS '25)*. Association for Computing Machinery, New York, NY, USA, 576–594. <https://doi.org/10.1145/3715336.3735786>
- [17] Jaziar Radianti, Tim A. Majchrzak, Jennifer Fromm, and Isabell Wohlgenannt. 2020. A systematic review of immersive virtual reality applications for higher education: Design elements, lessons learned, and research agenda. *Computers & Education* 147 (2020), 103778. <https://doi.org/10.1016/j.compedu.2019.103778>
- [18] Ryo Suzuki, Mar Gonzalez-Franco, Misha Sra, and David Lindlbauer. 2025. Everyday AR through AI-in-the-Loop. In *Proceedings of the Extended Abstracts of the CHI Conference on Human Factors in Computing Systems (CHI EA '25)*. Association for Computing Machinery, New York, NY, USA, Article 779, 5 pages. <https://doi.org/10.1145/3706599.3706741>
- [19] Mihail Terenti, Laura-Bianca Bilius, Andrei Braicov, Tatiana Veverița, Ilona Popovici, and Radu-Daniel Vatavu. 2026. Mapping Technological Dimensions of XR-based Cognitive Augmentation Systems. In *Proceedings of the 2026 International Conference on Development and Application Systems (DAS '26)*. 298–303. <https://doi.org/10.1109/DAS69882.2026.11553363>
- [20] Radu-Daniel Vatavu. 2026. Extended Reality Thinking: Why Interacting in XR Environments Calls for New Cognitive Skills. In *Proceedings of the Extended Abstracts of the 2026 CHI Conference on Human Factors in Computing Systems (CHI EA '26)*. Association for Computing Machinery, New York, NY, USA, Article 286, 6 pages. <https://doi.org/10.1145/3772363.3798484>
- [21] T. Veverița and A. Braicov. 2026. Computational thinking and critical thinking in the context of extended reality-based learning environments. In *Proceedings of ICVL 2026, the 21st International Conference on Virtual Learning*. 481–502. <https://doi.org/10.58503/icvl-v21y202641>
- [22] Bob G. Witmer and Michael J. Singer. 1998. Measuring Presence in Virtual Environments: A Presence Questionnaire. *Presence* 7, 3 (1998), 225–240. <https://doi.org/10.1162/105474698565686>
- [23] Ying Xu, Stacy Branham, Xinwei Deng, Penelope Collins, and Mark Warschauer. 2021. Are Current Voice Interfaces Designed to Support Children's Language Development?. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 633, 12 pages. <https://doi.org/10.1145/3411764.3445271>