Human Interaction and Emerging Technologies (IHIET 2024), Vol. 157, 2024, 297–305
https://doi.org/10.54941/ahfe1005489

Collaborative Learning Through XR. A Study of Eye- and Hand-Based XR Interactions to Support Collaborative Learning in the Chemistry Classroom

Wisanukorn Boribun and Frank Heidmann

University of Applied Science Potsdam, Brandenburg, 14469, Germany

ABSTRACT

In this study, we examine the utilization and efficacy of Extended Reality (XR) technologies in education. The focus is on understanding how XR interactions, particularly those centred around eye and hand movements, can enhance collaborative learning. To achieve this, we developed a specialized XR application tailored for this purpose. By leveraging augmented reality technologies, we created an interactive and immersive learning environment where fundamental chemical concepts can be visually and tangibly represented. The application enables students to interact with and study virtual sugar molecules, facilitating collaborative learning as they learn to distinguish between them. Additionally, the application incorporates eye-tracking and hand-tracking technologies, facilitating natural interaction and collaboration among students. Through comprehensive analysis and multiple application tests, we investigate how these advanced technologies can enhance understanding of basic chemical concepts and promote collaboration among students. Finally, we discuss the impact of such technologies on the overall learning environment and classroom dynamics, highlighting both the advantages and challenges of integrating XR technologies into teaching.

Keywords: XR (extended reality), VR (virtual reality), Collaborative learning, Augmented reality, Education, Chemistry classroom, Eye-tracking, Hand-tracking, Eye-based interaction

INTRODUCTION

The digital transformation is reshaping various sectors of our society, with education undergoing significant changes (Radianti et al., 2020). Amid this evolution, Extended Reality (XR), encompassing Virtual Reality (VR), Augmented Reality (AR), and Mixed Reality (MR), emerges as a powerful tool for fostering interaction and collaboration, especially within educational settings (Merchant et al., 2014; Billinghurst & Duenser, 2012; Dunleavy & Dede, 2013). The integration of XR and VR into education holds the potential to enrich learning experiences by facilitating collaboration and shared comprehension (Radu, 2014; Akçayır & Akçayır, 2017). Collaborative learning, in particular, assumes a pivotal role, enabling learners to collaborate on problem-solving, exchange ideas, and glean insights from each other (Johnson et al., 2014). This collaborative approach not only enhances understanding and knowledge absorption but also nurtures essential skills such as communication, teamwork, and critical thinking (Ott & Freina, 2015). Despite the immense promise of XR and VR, many applications remain segregated (Dalgarno & Lee, 2009). Shared experiences often rely on screen streaming or verbal explanations, potentially excluding individuals who are not directly engaged (Slater & Wilbur, 1997). Even in scenarios where multiple users utilize VR or XR headsets, shared experiences often remain confined to specific contexts, such as multiplayer VR games (Billinghurst et al., 2015). Addressing these challenges, we have developed a prototype harnessing the eye and hand tracking functionalities of the HoloLens 2. This prototype facilitates collaborative interactions with 3D elements within an XR environment. By tracking participants gazes within the XR scene, users can effortlessly identify the elements they are currently viewing. Consider, for instance, the ability to observe what your partner is currently focused on. This capability could significantly enhance communication and shared understanding, as each participant gains insight into the other's focus. Moreover, it promotes collaboration by enabling participants to concentrate on shared elements or issues, facilitating joint problem-solving (Yuen et al., 2011).

Our current prototype is tailored to explore the intricate molecular structures of sugar molecules within a collaborative XR environment (Santos et al., 2014). Teams of users can collaboratively delve into molecules such as fructose, glucose, galactose, and mannose, gaining deeper insights into their unique properties. To evaluate the functionality and cognitive impact of our XR prototype, we recruited 19 participants for assessment, ensuring they lacked prior knowledge or involvement in chemistry or related fields concerning sugar molecules.

Our work contributes to: (i) the exploration of XR technology's potential to enhance collaborative learning in educational contexts, (ii) the development of a novel prototype leveraging eye and hand-tracking capabilities to foster interactive learning experiences, and (iii) the provision of valuable insights into user experience and cognitive impacts, contributing to the broader discourse on XR in education (Wojciechowski & Cellary, 2013).

DEVELOPMENT AND STATUS OF XR IN EDUCATION

Extended Reality (XR) is increasingly gaining popularity as an educational technology. It is extensively used in educational institutions to facilitate asynchronous collaboration among learners at various locations (Merchant et al., 2014). A contemporary example of XR application in education is "ClassVR," which integrates 360-degree experiences and Virtual Reality (VR) into the learning process. While there are already various XR and VR applications that enable and promote collaboration, there remains a need for further development and research in many areas, especially in education. The goal should be to create and improve XR environments that optimally support collaborative learning. Furthermore, XR is being applied in various other fields such as medicine, real estate, retail, and entertainment. For instance, in medicine, XR allows for the simulation of surgeries to train medical professionals. In the real estate industry, potential buyers or tenants can take virtual tours of properties. In retail, customers can visualize and test products in an XR environment, and in the entertainment industry, XR enables immersive gaming and experiences.

Although the field of Virtual Reality (VR) is experiencing continuous innovations and advancements, the integration of eye-tracking technologies into many modern VR headsets is not yet fully implemented. This does not necessarily represent a deficiency but rather highlights unexploited potential for enhanced interaction mechanisms. With the use of eye-tracking technologies, users would no longer depend on aiming at an object in the XR environment with a controller to interact with it. Instead, they could select which element to interact with through eye movements. This could significantly increase the speed of interaction and improve the user experience. Therefore, there is an increasing need for further research and development in this area to integrate eye-tracking and other intuitive interaction methods into XR systems.

In our work, we explore new ways of visualization and interaction in XR systems to enhance collaborative learning within the XR environment. Unlike previous studies that focused on specific aspects of XR systems, our goal is to develop a system that provides a comprehensive and immediate overview of the shared learning environment (Billinghurst & Dünser, 2012).

COGNITIVE ASPECTS OF EYE-BASED XR INTERACTIONS

To explore the cognitive aspects of eye-based XR interactions, it is essential to have a solid understanding of eye-tracking technology and its integration into human-computer interfaces. Majaranta and Bulling (2014) provided a comprehensive overview of the development of eye-tracking technology, emphasizing its central role in medical and psychological research before its application in interactive, gaze-based applications. The implementation of eye-tracking in human-computer interaction relies on the eye's ability to convey cognitive processes and intentions (Majaranta & Bulling, 2014). The link between cognitive processes and eye movements is particularly significant for the design of XR systems. Majaranta and Bulling (2014) highlighted how eye tracking can enhance user interaction by leveraging natural eye movements, a principle applicable to XR applications for creating intuitive and efficient user interfaces.

The primary goal of our test prototype was to apply these principles and findings in practice. Specifically, we aimed to utilize natural eye movements to analyse user interactions in the XR environment. For instance, gaze fixation on an object, observed long enough to be processed by the brain's visual system, can serve as a cognitive task triggering interactions in virtual environments. By integrating eye-tracking technology into our prototype, we seek to gain deeper insights into the cognitive aspects of eye-based XR interactions.

FIXATION DURATION IN HUMAN-COMPUTER INTERACTION

Within the context of eye movement research, fixation is defined as a sustained visual focus on a specific point (Negi and Mitra, 2020). Longer fixations are believed to indicate that a person is processing or interacting with the information at that point. When developing our prototype, we considered findings from a study by Negi and Mitra (2020), who emphasized the importance of fixation duration in human-computer interaction. They found that longer fixation durations often indicate deeper cognitive involvement. Based on this understanding, we adjusted the interactivity of our user interface, designing it to activate only after a fixation duration of 300 ms and allowing a tolerance of up to 500 ms (Negi and Mitra, 2020). This design approach serves to improve the user experience in two important ways: firstly, by ensuring that interactions and information become accessible or visible only when the user is ready to engage with them, and secondly, by helping to prevent accidental selection of objects within the XR layer.

Table 1. Overview of fixation durations and their impact on learning processes based on findings by Negi and Mitra (2020).

Fixation Duration	Typical Significance	Potential Impact on Learning
Under 100 ms	Often too short for conscious perception	May indicate distraction or rapid scanning
100–200 ms	Brief fixation, limited information intake	Suggests minimal engagement or familiarity
200–300 ms	Normal for reading simple texts	Standard fixation duration for text learning
300–500 ms	Longer fixation, deeper cognitive engagement	Indicates intense engagement, likely comprehension
500–800 ms	Very long fixation, potential difficulty or interest	Could signify confusion or deep interest
Over 800 ms	Extremely long fixation uncommon in normal reading	May suggest cognitive overload or intense focus on a complex issue

XR PROTOTYPE (XVGAR)

The prototype is implemented using Unity 2021.3.5f1 and the Mixed Reality Toolkit (MRTK), designed specifically for standalone use with the HoloLens 2 mixed reality glasses, without needing a wired connection to a computer. A crucial feature of our prototype is gaze interaction, facilitated by the HoloLens 2's eye-tracking capability. This technology captures the user's gaze direction in real-time, accurately pinpointing their focus point. This is achieved through raycasting, where an invisible virtual beam, originating from the user's eyes, determines their gaze direction and collects collision data with 3D objects in the XR environment. When the raycast intersects with a 3D object, such as interfaces in the XR environment or a model of a sugar molecule, that object becomes the user's current focus point.

When using eye-tracking technologies, it is common to capture the gaze direction of each eye separately. However, since we naturally focus on a specific point with both eyes, determining a midpoint between the two gaze directions is necessary. This process, called interpolation, calculates a single central focus point from the gaze directions of both eyes. This method is particularly crucial in three-dimensional applications like the HoloLens 2, enabling precise interaction with elements in three-dimensional space. Additionally, eye-tracking systems account for small deviations, known as "vergence," to ensure accurate calculations and determine the precise midpoint between the two gaze directions.

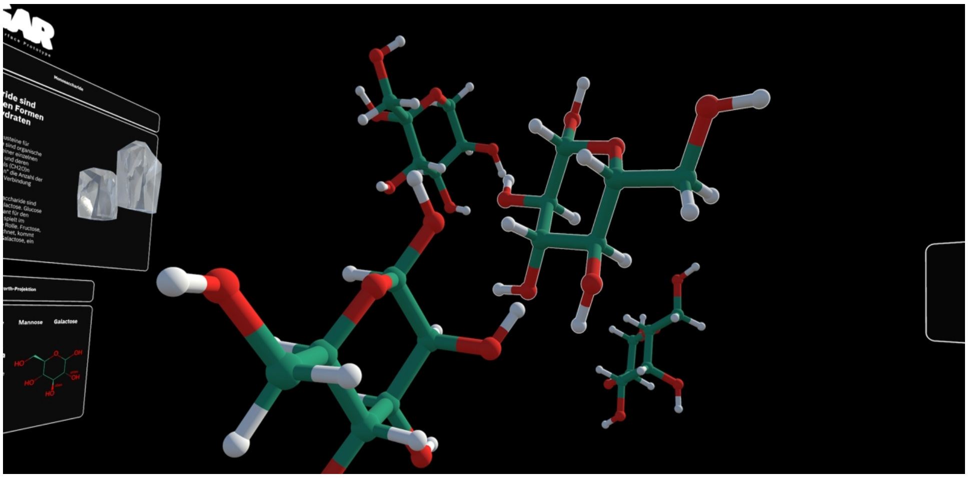
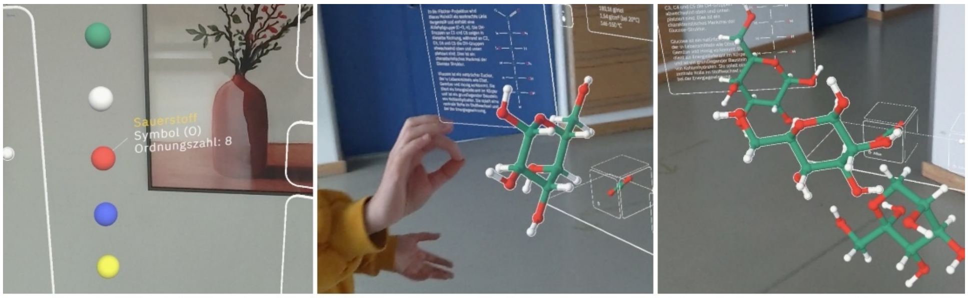


Figure 1: An exemplification of the XVGAR prototype. An eye-based interaction (fixation) is displayed, enabling users in the XR-Environment to identify the molecule currently being observed - glucose, fructose, mannose, and galactose.



a) b) c)

Figure 2: a) The graphic illustrates a tool-tip within the prototype, which becomes visible upon achieving a fixation duration of 300ms. Specifically, the sphere (oxygen atom) detects the user's fixation and subsequently displays pertinent information. b) This scenario incorporates eye-based interaction in conjunction with hand-tracking capabilities. The user employs a pinching gesture to interact with the object identified through the gaze-tracking mechanism. c) This graphic depicts the synchronization of fixation, indicating that all users are collectively focusing on a single molecule within the XR environment, as other sugar molecules are not highlighted.

The synchronization of the captured gaze information is done using the Photon Unity Networking (PUN2) plugin. PUN is a network plugin for Unity that facilitates real-time data exchange between users. In our project, we utilized PUN to synchronize the gaze information captured by the eye-tracking function. This gaze data, packaged into data packets, is transmitted to all connected users. Each packet contains detailed information about the position and orientation of the raycast, as well as the objects struck by the raycast. When a user fixates on a specific object, it becomes visible to the entire user group. After a designated fixation time (300–500ms), a visual signal is generated, indicating that the object is being viewed, enabling other users to precisely identify the object the first user is focused on.

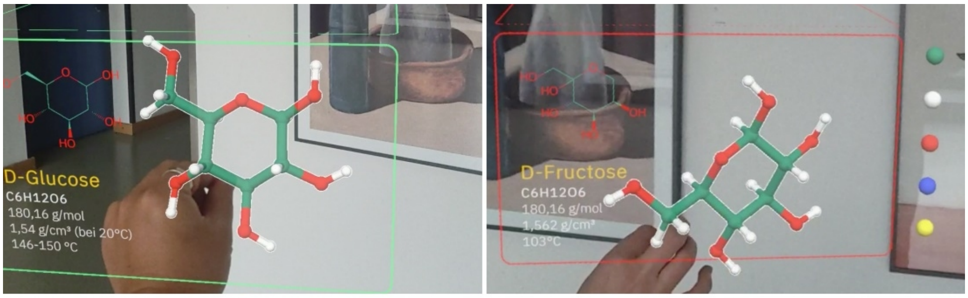
In addition to gaze data synchronization, we have also implemented hand tracking data synchronization. This means that with each shift of an object, data packets are transmitted, ensuring that every change is immediately visible on the displays of all users. This facilitates seamless and intuitive collaboration as users work on the same object in real time, with modifications visualized instantly. Technically, each interaction with the object generates a data packet sent via a network protocol to all connected devices. These packets contain specific information about the position and movement of the manipulated object, interpreted and implemented by other devices to maintain a cohesive representation of the interaction for all users.

In designing the prototype, we placed significant emphasis on interaction with the physical world. The interface is designed to be translucent, allowing users to perceive their physical environment while simultaneously interacting in augmented reality. This ensures that the interface does not impede physical tasks' performance and that users remain connected to their real environment. By integrating these technologies and design principles, we have developed a prototype that serves as an effective tool for collaborative work in augmented reality (Greenwald et al., 2017). These characteristics render the prototype ideal for our learning scenario and the execution of our study.

EVALUATION

In our initial formative study, we evaluated the effectiveness of the XR system, particularly focusing on its ability to simplify complex chemical concepts and promote collaborative learning. The study comprised two phases. In the first phase, a tutor explained the properties of existing sugar molecules, the meaning of different projection types, and how to distinguish and interpret them. In the second phase, participants had the opportunity to examine sugar molecules in the XR environment and collaboratively decide which sugar molecule should be inserted into each differentiation system.

We invited a group of nineteen participants, not directly involved in the field of chemistry, to test our XR prototype in small groups. Surprisingly, we could almost forego an explanation of the controls, as participants perceived eye control as an intuitive method for interaction in the XR environment. This improved the naturalness of the communication process and allowed users to quickly familiarize themselves with control and communication at the XR level. Using eye and hand control felt like a natural interaction for most.



a) b)

Figure 3: a) The XR environment features a synchronized system for distinguishing sugar molecules. In this illustration, the user places the correct molecule in the corresponding field. The box changes its color to green to visualize the correct placement of the molecule. b) This graphic displays the distinction system when an incorrect molecule is returned to the wrong field, causing the box to turn red, visible to all users.

Interestingly, our findings suggested that our XR system enabled participants to distinguish between different sugar molecules more quickly. This could be attributed to the interactive and playful way molecules were differentiated together, possibly due to the increased focus enabled by direct gaze control. Besides improving communication, our system also appeared to enhance individual learning and information processing abilities, indicating its potential to accelerate the learning process and enhance the overall user experience. Essentially, this study provided valuable insights into the potential of integrating eye-based interactions into XR systems for educational purposes. The results underscore the potential of XR technology to transform educational practices by creating engaging, intuitive, and collaborative learning environments.

CONCLUSION

In this work, we have introduced a comprehensive XR system based on eye and hand tracking interactions. We presented our investigations into the use of eye-tracking technologies in XR environments and described the development of our prototype employing these technologies. Moreover, we explored the cognitive aspects of eye-based XR interactions, highlighting the interplay of cognitive processes and eye movements. The possibility of utilizing eye-tracking technology with XR to synchronize fixations opens new opportunities for various fields, not just in educational contexts. Currently, we are testing our prototype in several studies, particularly aiming to enhance the user experience in collaborative XR environments. We are reflecting on the use and development of our technologies and presenting the insights gained. Our work aims to contribute to the continuous improvement of these tools and the resulting research in XR technologies. However, we believe that XR technologies should not replace traditional teaching methods and books, but rather complement them. They should serve to enhance students' collaboration and interaction, making learning more effective and engaging.

ACKNOWLEDGMENT

The authors would like to thank all participants and chemistry teachers for their participation in the evaluation. Their educational insights were invaluable to us.

REFERENCES

Akçayır, M. & Akçayır, G. (2017). Advantages and challenges associated with augmented reality for education: A systematic review of the literature. Educational Research Review, 20, 1–11. https://doi.org/10.1016/j.edurev.2016.11.002

Billinghurst, M., Clark, A. & Lee, G. (2015). A Survey of Augmented Reality. Foundations And Trends in Human-computer Interaction (Print), 8(2–3), 73–272. https://doi.org/10.1561/1100000049

Billinghurst, M. & Duenser, A. (2012). Augmented Reality in the Classroom. Computer, 45(7), 56–63. https://doi.org/10.1109/mc.2012.111

Dalgarno, B. & Lee, M. J. W. (2009). What are the learning affordances of 3-D virtual environments? British Journal Of Educational Technology, 41(1), 10–32. https://doi.org/10.1111/j.1467-8535.2009.01038.x

Dunleavy, M. & Dede, C. (2013). Augmented Reality Teaching and Learning. In Springer eBooks (S. 735–745). https://doi.org/10.1007/978-1-4614-3185-5_59

Greenwald, S., Corning, W. & Maes, P. (2017). Multi-User Framework for Collaboration and Co-Creation in Virtual Reality. MIT Media Lab. Abgerufen von https://dspace.mit.edu/handle/1721.1/108440

Johnson, D. W. & Johnson, R. T. (2014). Cooperative learning in 21st century. Journal Article, 30(3), 841–851. Abgerufen von https://doi.org/10.6018/analesps.30.3.201241

Majaranta, P. & Bulling, A. (2014). Eye Tracking and Eye-Based Human–Computer Interaction. In Human-computer interaction series (S. 39–65). https://doi.org/10.1007/978-1-4471-6392-3_3

Merchant, Z., Goetz, E. T., Cıfuentes, L., Keeney-Kennicutt, W. & Davis, T. J. (2014). Effectiveness of virtual reality-based instruction on students' learning outcomes in K-12 and higher education: A meta-analysis. Computers & Education, 70, 29–40. https://doi.org/10.1016/j.compedu.2013.07.033

Negi, S. & Mitra, R. (2020). Fixation duration and the learning process: an eye tracking study with subtitled videos. Journal Of Eye Movement Research, 13(6). https://doi.org/10.16910/jemr.13.6.1

Ott, M. & Freina, L. (2015). A Literature Review on Immersive Virtual Reality in Education: State Of The Art and Perspectives. eLearning And Software For Education. https://doi.org/10.12753/2066-026x-15-020

Radianti, J., Majchrzak, T. A., Fromm, J. & Wohlgenannt, I. (2020). A systematic review of immersive virtual reality applications for higher education: Design elements, lessons learned, and research agenda. Computers And Education (Print), 147, 103778. https://doi.org/10.1016/j.compedu.2019.103778

Radu, I. (2014). Augmented reality in education: A meta-review and cross-media analysis. Personal And Ubiquitous Computing, 18(6), 1533–1543. https://doi.org/10.1007/s00779-013-0747-y

Santos, M. E. C., Chen, A., Taketomi, T., Yamamoto, G., Miyazaki, J. & Kato, H. (2014). Augmented Reality Learning Experiences: Survey of Prototype Design and Evaluation. IEEE Transactions On Learning Technologies, 7(1), 38–56. https://doi.org/10.1109/tlt.2013.37

Slater, M. & Wilbur, S. (1997). A Framework for Immersive Virtual Environments (FIVE): Speculations on the Role of Presence in Virtual Environments. Presence (Cambridge, Mass.), 6(6), 603–616. https://doi.org/10.1162/pres.1997.6.6.603

Wojciechowski, R. & Cellary, W. (2013). Evaluation of learners' attitude toward learning in ARIES augmented reality environments. Computers And Education (Print), 68, 570–585. https://doi.org/10.1016/j.compedu.2013.02.014

Yuen, S. C., Yaoyuneyong, G. & Johnson, E. (2011). Augmented Reality: An Overview and Five Directions for AR in Education. Journal Of Educational Technology Development And Exchange (Print), 4(1). https://doi.org/10.18785/jetde.0401.10