2025 IEEE Conference Virtual Reality and 3D User Interfaces (VR) | 979-8-3315-3645-9/25/$31.00 ©2025 IEEE | DOI: 10.1109/VR59515.2025.00061

XRXL: A System for Immersive Visualization in Large Lectures

Kabir Batra* Purdue University
Zirui Zhang† Purdue University
Shuwen Yang‡ Purdue University
Anima Agrawal§ Purdue University
Yiyin Gu¶ Purdue University
Bedrich Benes‖ Purdue University
Alejandra Magana** Purdue University
Voicu Popescu†† Purdue University

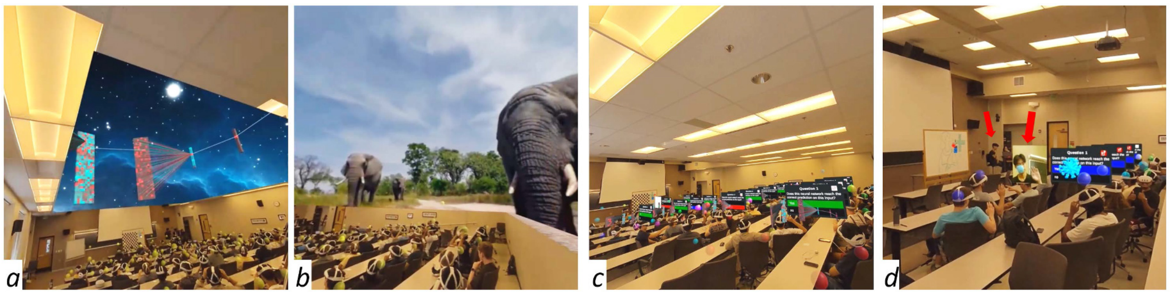

Figure 1: *XRXL* deployed in a large lecture with 82 students. The images are frames acquired with an XR headset from the back of the classroom, showing what the students see. The instructor has virtually retracted the ceiling to make room for a large-scale 3D visualization of a neural network (*a*). The instructor has turned the classroom into a 360° theater to take the class on an African safari (*b*). The instructor has partitioned the class into groups of four students, each with their own neural network 3D visualization and 2D panel for answering questions (*c*). The instructor pays a virtual visit to a group, without leaving the instructor's desk (*d*).

Abstract

This paper describes *XRXL*, an extended-reality system for increasing student engagement in large lectures. Students wear XR headsets to see 3D visualizations controlled by the instructor. The instructor can virtually retract the roof and walls of the classroom to allow for large-scale visualizations that extend beyond the physical boundaries of the classroom, or to turn the classroom into a 360° theater. The instructor can also partition the classroom into small groups of students and to assist individual groups as needed. *XRXL* was tested in an IRB-approved user study with 82 students in the context of a mock-lecture on neural networks. To the best of our knowledge, the study is the largest deployment of a co-located collaborative XR application to date. The study shows that students had a favorable opinion of *XRXL*, that *XRXL* had a low task load, an acceptable usability level, and that it did not cause cybersickness.

Index Terms: Computing methodologies—Computer graphics—Graphics systems and interfaces—Mixed / augmented reality; Applied computing—Education—Interactive learning environments

1 Introduction

Despite their best efforts, colleges and universities struggle to keep class sizes down. For popular majors such as computer science, data science, or engineering, many undergraduate courses hold lecture with 50, 100, or even several hundreds of students. Such large lectures are often reduced to an instructor monologue accompanied by slides projected on a screen at the front of the classroom. Students cannot interact with the instructor, with each other, or with the slides; they lose interest, and they disengage. Campuses do experiment with classrooms that are reconfigurable from an instructor presentation mode to a student group mode, but these classrooms require specialized infrastructure, such as pivoting desks and retractable displays, which are expensive to acquire and maintain.

The potential of immersive visualization to engage students has long been known. A student wearing a virtual reality (VR) headset sees visualizations in 3D, with appropriate depth perception; they can select the desired view intuitively, by moving their head; and they can interact with the visualization through physical motions that scaffold learning, as shown by embodied cognition research. However, immersive visualization has been primarily reserved for individual or small group learning in virtual laboratory or home settings. One reason is that until recently VR headsets were not portable, requiring nearby workstations to assist with rendering, and nearby sensors for outside-looking-in tracking. Another reason is that the VR headset isolates the student from their physical surroundings, preventing them from seeing the instructor, their laptop, or their fellow students. Finally, high technology costs made that one or a few headsets had to be shared by multiple students. We put forth that technology has reached a stage where immersive visualization is becoming tractable in the context of large lectures. We now have completely untethered extended reality (XR) headsets (e.g., Meta's Quest 3 [53]), with on-board power, rendering, and tracking, and with a passthrough mode that allows the learner to see the important elements of their physical surroundings.

In this paper we present *XRXL* (eXtended Reality eXtra Large), a novel system for immersive visualization in large lectures. *XRXL* is illustrated in Fig. 1, as well as in the video accompanying our paper. The instructor controls *XRXL* from their laptop or classroom computer. Students wear XR headsets in passthrough mode to see the physical classroom augmented with 3D visualizations.

In *instructor presentation mode*, students see a large-scale 3D visualization controlled by the instructor. As the instructor points at the visualization on their laptop, students see where the instructor

*e-mail: batra14@purdue.edu
†e-mail: zhan4192@purdue.edu
‡e-mail: yang1949@purdue.edu
§e-mail: agraw208@purdue.edu
¶e-mail: gu251@purdue.edu
‖e-mail: bbenes@purdue.edu
**e-mail: admagana@purdue.edu
††e-mail: popescu@purdue.edu

2642-5254/25/$31.00 ©2025 IEEE
DOI 10.1109/VR59515.2025.00061

Authorized licensed use limited to: Purdue University. Downloaded on April 26,2025 at 15:13:31 UTC from IEEE Xplore. Restrictions apply.

is pointing through a virtual laser beam that moves in sync with the laptop cursor. The instructor can virtually retract the roof and walls of the classroom to make room for a 3D visualization that extends beyond the physical classroom (Fig. 1, *a*), or to turn the classroom into a 360° theater (Fig. 1, *b*). At the press of a button on the *XRXL* interface, the instructor can switch to *student group mode*.

In *student group mode* the classroom is virtually partitioned into small groups of students, e.g., four students in Fig. 1, *c*. Each group has their local 3D visualization, as well as a virtual 2D display, for example to consult lecture notes, or to answer questions. Each of the students in the group has a sphere of the same color floating above their head, for students to know who is in their group. A group can ask for help from the instructor, who can pay the group a virtual visit, without leaving the instructor's desk (Fig. 1, *d*). During such a visit, a live video sprite of the instructor appears by the group's 2D display; furthermore, the instructor can see the group and their visualization from the vantage point of any of the students in the group, e.g., from the vantage point of the student asking the question, such that the instructor can quickly gain familiarity with the context of the question and provide adequate assistance.

We have evaluated *XRXL* in a user study (N = 82) approved by Purdue University's Institutional Review Board (IRB-2024-119). The study aimed to answer two overarching research questions: **RQ1**–can *XRXL* provide immersive 3D visualization to the students of a large lecture, as they follow the instructor presentation and as they work in small groups?, and **RQ2**–does *XRXL* perform adequately in terms of usability, task load, cybersickness, and subjective student opinion? The participants were undergraduate students (90%) and they used *XRXL* in the context of a mock-lecture on the architecture and functioning of neural networks. To the best of our knowledge, our study is the largest deployment to date of a co-located collaborative XR application. Participants experienced all *XRXL* features. The study shows that students enjoyed seeing 3D visualizations (4.5/5.0), that they liked that the classroom can turn into a 360° theater (4.5/5.0), that they liked that a student group can get the instructor's help (4.4/5.0), and that they thought *XRXL* can make large lectures less boring and more engaging (4.3/5.0). The study also shows that the students were slightly negative or neutral regarding having to wear the XR headset (2.7/5.0), regarding not being able to see each others' faces (3.1/5.0), and regarding being distracted from lecture (2.8/5.0). Students declared themselves reluctant to purchase a $500 headset (2.2/5.0), expecting that the headset be provided to them (3.9/5.0). Finally, *XRXL* received a "Good/OK" usability score, i.e., an SUS score of 65.9, a low cybersickness score, i.e., a total SSQ score of 14.0, and a low task load score, i.e., a TLX overall raw score of 2.9/10.0.

2 Prior Work

We first give a brief overview of educational science research that confirms the need and informs the design of *XRXL* (Sec. 2.1), and we then discuss prior work on collaborative XR (Sec. 2.2), with an emphasis on large-scale XR application deployments (Sec. 2.3).

2.1 Large Lecture Effectiveness

Education science research has long signaled that large lectures have limited effectiveness [60, 37], pointing to the fact that they fail to engage students [30, 65]. Reconfigurable classrooms, in which the desks and chairs can be rearranged to alternate instructor presentation with student group work, support active learning and teaching approaches [28, 27] that result in higher levels of student engagement [10, 49, 21]. One challenge is that physically reconfigurable classrooms require expensive furniture. Even the best equipped campuses have only a few such classrooms. Another challenge is that a class of 80 students yields 20 groups of four students, and the instructor cannot easily move from group to group to monitor progress and provide assistance to individual groups as needed. Motivated by the success of physically reconfigurable classrooms, *XRXL* leverages XR technology to allow for the *virtual* reconfiguration of the classroom. An XR headset allows instantiating 2D and 3D displays anywhere, as needed to partition the classroom with any granularity, without the expense of physical displays that have to be deployed and retracted to switch from instructor presentation to student group work, and then back. With XR, *all* classrooms on campus can become reconfigurable.

In addition to flexible and inexpensive classroom reconfiguration, XR also brings the benefit of supporting learning through embodied cognition [18]. XR increases the level of sensorial and bodily engagement, which translates to knowledge retention [31]. Seeing visualizations and simulations does improve learning [14], as do even simple hand manipulations or gestures [13, 71]. Bodily engagement can lead to cognitive engagement [39] and ultimately to the productive engagement of students [5]. Student engagement benefits beyond academic achievement, and it has been shown to contribute to cognitive and social development [45, 23, 35].

2.2 Collaborative XR

XRXL affords collaboration between an instructor and tens of students collocated in the same physical classroom, so we first discuss prior work on collocated collaboration with the help of XR. Furthermore, the large number of students precludes that the instructor visit each of the tens of student groups by physically walking through the classroom to join groups one at a time; therefore, *XRXL* calls for the instructor to be able to pay "remote" visits to the student groups, without having to leave their desk, virtually crisscrossing the large classroom to assist each student group in need; as such, we also discuss prior work on remote collaboration aided by XR.

Collocated collaborative XR. In co-located collaboration the users share the same space [70]. One challenge is an accurate or at least consistent virtual to real alignment for the multiple users [46]. Another challenge is to orchestrate the shared use of the same physical space, which has been addressed, for example, by the virtual replication of the high-contention regions of the real-world environment [75]. In the context of our *XRXL* system, the alignment problem is exacerbated by the large size of the lecture room, and the space contention problem is simplified by the fact that our students are seated or standing in place in the classroom, avoiding the problem of tens of headset wearers moving about in a shared space.

Remote collaborative XR. XR has been leveraged in remote collaboration, where it allows collaborators to see each other as well as the shared workspace. One challenge is real-time high-fidelity acquisition of the collaborators and of the workspace, which can be dynamic and have intricate geometry and color [69]. Providing the user with a free viewpoint visualization of their remote collaborators requires depth acquisition [2], while video only acquisition has simplicity and robustness advantages [7]. An important concern in XR remote collaboration is the degree to which the collaborators feel as if they are all present in the same room, i.e., their sense of co-presence [33, 66]. We refer the reader to a comprehensive review of prior work on XR-enabled remote collaboration [72].

The remote collaboration needs of *XRXL* are limited to allowing the instructor to visit student groups, which are addressed satisfactorily by modeling the instructor with a real-time video sprite. For the instructor to see a student group, *XRXL* relies on first-person video feeds acquired by student headsets, which require stabilization. Stabilization implies re-rendering from a stable virtual viewpoint, which requires real-time depth acquisition from multiple viewpoints. Depth acquisition can be bypassed through approximations such as warping [40, 42] or homographies [11, 38], aided by optical flow [74], correspondences [64], or additional sensors [59]. In our context, the video is acquired by the headset, which provides each frame's extrinsic camera parameters, and we use a homography for its attractive performance to computational cost ratio.

Authorized licensed use limited to: Purdue University. Downloaded on April 26,2025 at 15:13:31 UTC from IEEE Xplore. Restrictions apply.

2.3 Large-Scale XR Deployments

Most XR applications are designed for individual or few users, for example to provide assistance to technicians [24, 17, 16] or surgeons [47, 41, 67] during training or on the job, or to help students learn [68, 36, 58]. The need, opportunities, and challenges of scalable extended reality have been documented in future research agendas [48], which note the importance of scaling with the number of users, at the same time pointing out that the challenges are best addressed through joint efforts that cross between disciplines such as computer graphics, computer vision, human-computer interaction, computer networking, and education science.

The first successes in deploying XR applications at scale took advantage of the proliferation of phones that implement a transparent display XR interface. Pokémon Go‘s success [15] transcended entertainment to increase motivation for physical exercise [8] and for learning English [73], and to help restaurants find more customers [62]. Stadium sport spectators can watch pitch-aligned replays on their phone [43]. In education, students use their phones to enhance learning, e.g., in chemistry [20]. *XRXL* provides *immersive* XR visualization to students, through an XR headset, which has the advantage of a larger field of view, of depth perception, and of not having to hold the phone. The advantage of phones is that they are already mass-deployed. Recent XR headsets [4] provide some degree of bidirectional video see-through capability, allowing headset wearers to see each others‘ faces. Headset inpainting, i.e., a diminished reality approach for removing the headset, in software, is an active research field with rapidly improving quality and frame rates [22, 25], which we will integrate in future *XRXL* prototypes.

XR has been used to make presentations more personal for individual audience members [63]. The level of audience engagement increases even if the real world experience is only augmented in the audio channel [61]. XR animated infographics have been shown to enhance communication effectiveness in business presentations [19]. Immersive XR has been used in large lectures in the context of synchronous distance education to implement hybrid classrooms with local and remote students; the instructor wears an XR headset that renders remote student video sprites in the empty seats of the classroom [29]. Whereas in the hybrid classroom the instructor is the only one wearing an XR headset, *XRXL* scales immersive visualization to reach each student of a large lecture. Sport spectators can enjoy the stadium experience from home, with an immersive viewing of a 360° video feed [44].

Untethered XR headsets that enable the large-scale deployment of immersive XR visualization have only recently become available, and their performance/price ratio continues to improve, e.g., Meta‘s Quest 3 [53] (released in October 2023) has higher resolution passthrough than the Quest Pro [55] (released in October 2022), at half the price. Another year later, i.e., in October 2024, the release of Quest 3S [54] lowers the *XRXL* cost per student from \$500 to \$300. The XR software stack is also maturing, providing support for scene proxy acquisition, for virtual to real alignment, for hand tracking, and for multi-user management, to the benefit of XR application developers and users. Leveraging these prior research results and technological advances, building and testing *XRXL* in a large lecture with 82 students is now possible.

3 System Design

We have set out to begin addressing the student engagement problem in large lectures with the help of XR. We first provide the system design rationale (Sec. 3.1), motivating the desired functionality, and then we discuss high-level system design implementation choices made to provide the desired functionality (Sec. 3.2).

3.1 Design Rationale

We aim to engage students in large lectures by leveraging XR (1) to provide students with immersive 3D visualization, and (2) to allow interleaving lecture segments where students follow the instructor‘s presentation with lecture segments where students work in small groups. We use the term immersive 3D visualization to denote visualization (a) that exhibits left-right eye disparity supporting depth perception, and (b) that allows the user to select the desired view naturally, by moving their head. To support immersive 3D visualization and small group collaboration in a context of a large lectures, *XRXL* was designed to provide the following functionality.

Immersive 3D visualization. Students should be able to see 3D visualizations, with depth cues, both stationary and dynamic, anchored in a shared 3D space. Students should be able to select the desired view naturally, by moving their head. Immersive 3D visualization promises to engage students beyond the 2D visualizations projected in traditional classrooms, through embodied cognition.

Unlimited display volume. The visualizations should not be confined to the boundaries of the physical classroom. Instead, the ceiling and walls of the classroom should be virtually retractable to make room for visualizations of any size and to turn the classroom into a virtual 360° theater. In addition to an improved spatial understanding of complex topics and 3D structures, like neural networks, such large-scale and surrounding visualizations promise readability, eloquence, and immersion, in support of student engagement. Allowing for visualizations to be displayed anywhere in the space surrounding the students, and at any scale, allows all students in the large classroom to see the visualization well, including those seated at the back of the classroom. Furthermore, presenting a visualization in a size commensurate to the size of the group of students allows building esprit de corps, making students feel part of the same, albeit large, group. For example, a large-scale visualization floating above the group of students has front-row students turn around, becoming aware of the students seated at the back, alleviating the student isolation common in large lectures.

Physical surroundings visualization. Students should be able to see the actual physical classroom, including the instructor and their fellow students. Seeing the physical classroom promises to reduce student isolation. Furthermore, compared to VR, allowing students to see their physical surroundings reduces the risk of cybersickness, as the passthrough background is updated at constant and high frame rate, independent of the graphics rendering load that could cause cybersickness-inducing frame rate fluctuations. Finally, allowing students to see their laptops enables efficient note taking with a conventional keyboard interface, as needed, for example, to complete a report for an individual or small group learning activity.

Student group mode. Students should be able to work in small groups. Switching between instructor and student group mode should be easy. Interleaving instructor presentation with working in small student groups promises to engage students.

Instructor assistance to student groups. A student group should be able to request assistance from the instructor, and the instructor should be able to pay a visit to any group. In order to make visiting any and many groups tractable, the instructor visits should be virtual, i.e., from the instructor‘s desk, without physical locomotion through the classroom. During the visit, the students of the group and the instructor should be able to see and hear each other, and the instructor should be able to assume the vantage point of any of the students in the group. These virtual visits promise to increase the efficiency with which the instructor can help the many groups of a large lecture, avoiding the delay and disruption of having to actually walk from one group to the next.

Adequate system usability. The system should provide an intuitive interface, the system should be comfortable to use in terms of physical encumbrance and cybersickness, and the system should function robustly, in order to allow for extended use.

Scalability. The system should be scalable with the number of students, to support large lectures, and with the number of classrooms, to support deployment across campuses. Scalability implies

Authorized licensed use limited to: Purdue University. Downloaded on April 26,2025 at 15:13:31 UTC from IEEE Xplore. Restrictions apply.

maintaining usability, but also controlling equipment, support, and logistics cost as the number of students and classrooms increases.

3.2 Design Implementation Choices

We now discuss the high-level *XRXL* implementation choices made to meet the desired functionality described above.

Immersive 3D visualization technology. One option for providing immersive 3D visualization is to upgrade the classroom projection system to a 3D projection system that alternates left and right eye images, which the students see using simple glasses that prevent an eye to see the image meant for the other eye, like in commercial 3D movie theaters. One shortcoming of this approach is that students cannot benefit from 3D visualization when working in groups. Furthermore, the classroom cannot be turned into a 360° theater without a substantial infrastructure investment to extend the screen and projector system all around the classroom, which precludes deployment to all campus classrooms.

Another option is to deliver the 3D visualization through XR headsets worn by each student. The important benefit is that the display is virtual, so the visualization can be placed anywhere and it can be of any size, without any additional cost. The headsets can be registered in the same 3D space so students see the 3D visualization at the same location relative to them. Delivering the 3D visualization through headsets allows partitioning the classroom into groups with any granularity, with each group seeing their own visualization. The challenges of the approach are that it requires an XR headset per student, and that XR headsets might not yet have the form factor and the technological maturity prerequisite for comfortable and robust use over extended periods of time.

We have selected the XR headset option as it supports both the instructor and the student group modes, because there are already XR headsets commercially available at price/performance points sufficient for prototype implementations, and because XR headsets are likely to advance in the near future. The optimism regarding XR technology is justified. Meta‘s Quest 3S [54], released after our study, allows running *XRXL* at the cost of $300 per student. Although the initial motivation of Meta‘s headsets was to power a metaverse closer to the virtual end of the VR/AR continuum, the emphasis has shifted to supporting applications that allow the user to see their physical surroundings. The same motivation convinced another trillion-dollar company to enter the XR space, supporting “spatial computing”, albeit currently at a prohibitive cost [4].

XR headset technology. An XR headset has to integrate the 3D visualization into the user‘s physical surroundings, which is a significant technological challenge. One approach is to allow the user to see the real world directly, with their own eyes, through a transparent surface. Such optical see-through headsets (e.g., Microsoft‘s HoloLens2 [57]) have the advantage that they provide the user with the best possible view of their physical surroundings. However, optical see-through headsets have the disadvantage of a small active field of view, i.e., the 3D visualization is limited to a small field of view within the user‘s natural field of view. Another disadvantage is that the visualization is transparent, which reduces visualization clarity, especially for bright backgrounds.

A second type of XR headset acquires the user‘s physical surroundings with one video camera per eye and displays the live feeds for each eye. The advantages of such video see-through headsets (e.g., Apple‘s VisionPro [4] and Meta‘s Quest 3 [53] and Quest 3S [54]) are a large active field of view, as well as support for true visualization opacity, allowing the visualization to completely erase the background, no matter its brightness. The disadvantage is that the user does not see the real world with their own eyes, but rather a live video feed of it, which comes at the cost of a lower field of view, dynamic range, and resolution. Another disadvantage is the unavoidable offset between the cameras and the user‘s eyes, which is particularly problematic for nearby objects.

Whereas optical see-through might be the ultimate goal of XR headset technology, overcoming its fundamental field of view and visualization opacity limitations has proven to be difficult. These limitations are particularly taxing in the context of our application. The limited active field of view would prevent a student from seeing a large 3D visualization in its entirety, and the student would have to scan it piece by piece by panning and tilting their view direction. The lack of support for visualization opacity precludes erasing the classroom ceiling and walls convincingly, as needed to accommodate large size or 360° visualizations. Video see-through is the approach of choice for current commercial headsets. Based on these considerations, we have opted for video see-through headsets for our prototype *XRXL* implementation. Should breakthroughs remove the current limitations of optical see-through, *XRXL* is ready to employ the best available XR headset.

Remote instructor/student group communication. For the instructor to be able to provide assistance to individual student groups efficiently, the instructor and a student group have to be able to communicate remotely. Remote collaboration is a classical application of XR. Approaches for XR-enabled remote collaboration can be classified on a continuum, based on the fidelity with which the remote party is acquired and rendered. At the low-fidelity end of the spectrum, the remote party is acquired with a video stream, which is sent to the local site where it is rendered on a virtual 2D screen. At the high-fidelity end of the spectrum, the remote party is acquired in 3D, and the 3D data is rendered at the local site from the local party‘s viewpoint. High-fidelity real-time 3D acquisition of humans remains challenging, especially without specialized equipment such as depth cameras and outside-looking-in acquisition rigs.

Based on these considerations we have opted to forgo depth acquisition and to capture the instructor and the group with a video stream alone. The instructor is captured with their computer‘s webcam, and the video feed is sent to the group asking for help, where it is displayed as a floating video sprite. To bypass the need for additional cameras, the student group is acquired using one of the eye cameras of the XR headset of one of the students. However, this first-person video is inadequate for providing the instructor with situational awareness, as the view direction swings unpredictably, often, and substantially, in sync with the student‘s head motions. Before the headset-acquired video stream can be shown to the instructor, it has to be stabilized. We stabilize the video using a plane at infinity 3D homography, which provides good results, bypassing the expense of depth acquisition.

User interface design. XR headsets provide a natural interface for selecting the desired view with six degrees of freedom through head motions and view direction rotations, a task that is challenging with a conventional mouse or keyboard interface. However, it is also the case that tasks that are relatively straightforward with a conventional interface, such as navigating nested menus or selecting small objects, can be challenging when wearing an XR headset, especially for a novice user. In order for *XRXL* to be eventually adopted in all classrooms across the campus, one cannot and should not expect that the instructor is an experienced XR user, nor that they have the time or interest to become one. Furthermore, having the instructor wear a headset hides their face from the students, precluding eye contact and hindering communication. Based on these considerations, we have designed the *XRXL* interface to allow the instructor to control the 3D visualization catered to students *without* having to wear an XR headset. Instead, the user controls the 3D visualization through the familiar interface of their computer. The instructor points at various parts of the visualization using their computer‘s mouse, which controls the virtual laser ray pointing at the 3D visualization catered to students.

Regarding the student interface, one option is to rely on handheld controllers that cast virtual rays for selection and buttons for clicking, and another option is to rely on hand tracking and gesture

Authorized licensed use limited to: Purdue University. Downloaded on April 26,2025 at 15:13:31 UTC from IEEE Xplore. Restrictions apply.

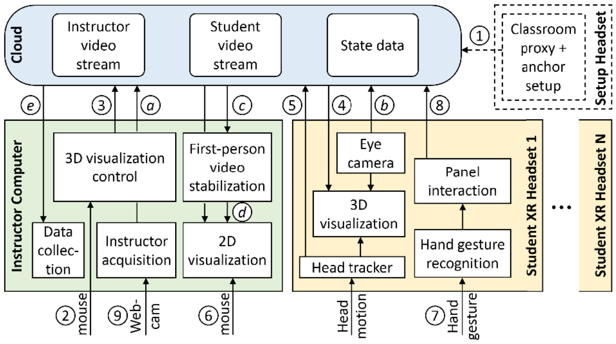

Figure 2: *XRXL* system architecture. Each of N students wears an XR headset to see 3D visualizations integrated into the classroom. The instructor controls the 3D visualizations through their computer, communicating with the headsets through a cloud server.

recognition. Controllers have the advantages of intuitiveness and robustness, and the disadvantages of perennially low batteries, and of tripling the number of devices that have to be managed. Gesture recognition avoids these disadvantages but it requires learning gestures, which can be misinterpreted, leading to user frustration.

Scalability. Rendering load is not a concern in terms of scalability with the number of students, as each headset renders their own and only their own visualization, as does the instructor computer. Networking load has to be managed by pre-loading all visualization objects on each headset. This way, when switching visualization mode or visualization object, the only data transfer needed is to communicate the id and not the description of the objects to be rendered. The highest network traffic is caused by the live communication between the instructor and a student group needing assistance, but this only entails bidirectional audio/video communication, which is well within the capabilities of today‘s networks. By relying on XR headsets, *XRXL* does not require any additional classroom infrastructure–the instructor can run the system from the classroom computer or from their own laptop. If each student owns an XR headset, *XRXL* can be deployed to all campus classrooms with little additional technology cost. We have already discussed the likely continued increase in performance/cost ratio of XR headsets. The technology acquisition cost is only one aspect of the cost of deployment to many students, many classrooms, and many campuses. Since XR technology can only be effective if adequate XR content is available, another aspect of the cost is content creation, which remains difficult. The present paper does not address the XR content creation bottleneck. Finally, the logistics of supporting a large number of headsets, and of setting up an XR large lecture, are daunting, and our pioneering study helps clarify where setup efficiency improvements are needed, and how, as discussed later.

4 XRXL System Architecture and Implementation

In this section we give an overview of the *XRXL* architecture (Sec. 4.1), we describe 3D visualization control by the instructor (Sec. 4.2), we describe the extension of the 3D visualization beyond the physical boundaries of the classroom (Sec. 4.3), and we describe the stabilization of a first-person student video to provide the instructor with an effective visualization of a student group (Sec. 4.4).

4.1 System overview

Fig. 2 gives an overview of the *XRXL* architecture.

Before lecture, during a setup phase, the instructor or a teaching assistant wears an XR headset and uses a handheld controller to acquire a geometric proxy of the classroom (1 in Fig. 2). The proxy defines the classroom floor, ceiling, walls, and rows of desks. The classroom boundary is defined by pointing with a virtual laser at the floor and ceiling corners of the walls, which we implement using Meta‘s Scene Model API [56]. The row desktop is defined by placing the handheld controller at the desktop corners. Also during the setup phase, the assistant defines a spatial anchor at the front of the classroom, which is then used by all student headsets to align their virtual coordinate system to the physical classroom.

During lecture, the instructor controls the 3D visualization using the 2D pointing device of their computer, i.e., the computer‘s touchscreen, touchpad, or mouse (2 in Fig. 2), to manipulate a conventional graphical user interface (GUI). 3D visualization control defines the objects that have to be rendered by each headset and updates the visualization state data on the cloud (3 in Fig. 2). The headsets poll the state data from the cloud (4 in Fig. 2) to update their 3D visualization. Each headset updates its position and orientation continually on the cloud (5 in Fig. 2).

In instructor mode, all headsets render the same large format 3D visualization. The instructor sees the visualization in a window on their computer screen and manipulates it with the mouse (6 in Fig. 2), for example to rotate it, to scale it, or to move it from the front to center of the classroom. The instructor 2D pointer is echoed by a 3D virtual laser pointer for the students to see where the instructor is pointing (Sec. 4.2). The instructor can virtually retract the classroom ceiling and walls to accommodate large-scale 3D visualizations or to turn the classroom into a 360° theater (Sec. 4.3).

In student group mode each headset renders the group‘s 3D visualization, as well as a 2D panel (Fig. 1, *c*). The visualization is placed above the desk in between the students of the group. The panel is vertical and at a fixed distance to the side of the group. A headset also renders a sphere above the head of each student in the group, for a student to easily identify their group mates. Students can interact with the panel through hand gestures (7 in Fig. 2) to answer multiple choice questions or to request help from the instructor, interactions transmitted through the cloud (8 in Fig. 2).

When the instructor assists a group, the webcam feed (9 in Fig. 2) is transmitted through a cloud video streaming service (*a* in Fig. 2). The instructor video stream is rendered for each student in the group on a rectangle next to the group‘s 2D panel (Fig. 1, *c*). The instructor can assume the viewpoint of any student in the group. The video feed from one of the eye cameras of the student headset is uploaded to the cloud (*b* in Fig. 2), downloaded by the instructor computer, stabilized (Sec. 4.4), and displayed (*d* in Fig. 2).

The computer and headsets communicate through several cloud services. We use the Photon Unity Networking (PUN) cloud service [1] for the state data, Meta developer hub casting service [50] for the student feed, and Agora‘s video streaming service [3] for the instructor feed. The instructor computer continually reads cloud data to maintain classroom awareness (*e* in Fig. 2).

4.2 3D Visualization Control

The instructor controls the 3D visualization catered to students using the conventional mouse + screen interface of their computer. One task is to replicate the instructor‘s 2D mouse pointer as a 3D laser pointer that shows students where the instructor is pointing (Fig. 3). This is done by computing the two 3D points A and B that define the laser ray segment. A is on the surface of the 3D object visualized, at the location where the instructor is pointing. A is computed by intersecting the instructor eye ray r with the 3D object. r is the visualization camera ray through the pixel covered by the instructor 2D pointer. B is placed at a distance d away from A along the ray to V. d depends on the visualization mode, with larger values for the instructor mode (e.g., 1 m), and smaller values for student group mode (e.g., 0.3 m).

Another task of 3D visualization control is to update on the cloud the state data that encodes the transform of the object with which the instructor interacts, for any change to be reflected in all student headsets. A third task of 3D visualization control is to give the instructor an overview of the current classroom configuration.

Authorized licensed use limited to: Purdue University. Downloaded on April 26,2025 at 15:13:31 UTC from IEEE Xplore. Restrictions apply.

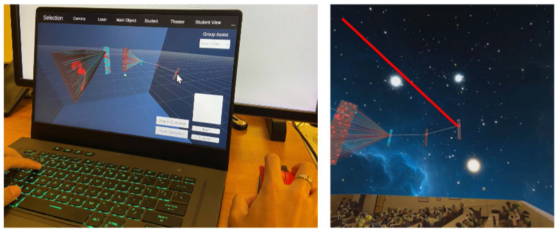

Figure 3: The instructor is pointing at the visualization using their laptop mouse (left), and students see in their headset a virtual laser pointer where the instructor is pointing (right).

We use a top-down classroom visualization that shows with colored dots the live student head positions read from the cloud state data. Students in the same group are shown with the same color.

4.3 Boundless 3D Display Volume

An important strength of XR is that it allows for large display volumes that do not have to be confined to the boundaries of the physical space. However, simply showing a large-scale visualization that goes beyond the boundaries of the physical space creates a confusing effect with the visualization intersecting the physical boundary. For example, if a 3D visualization extends beyond the ceiling of the classroom, the user will be confused by the conflicting depth cues of the closer ceiling and the farther 3D visualization. To make the visualization easier to parse, we place a virtual hole in the ceiling through which the 3D visualization protrudes.

The virtual hole is implemented by manipulating the ceiling geometry on which the video see-through is rendered. We model the ceiling with four rectangles. When the ceiling is fully closed, the four rectangles cover the entire ceiling. As the four rectangles retract, they leave a growing rectangular hole of the same aspect ratio as the ceiling. The passthrough video is not rendered over the hole, which effectively deletes the corresponding part of the ceiling. The user sees through the hole the skybox and the 3D visualization that appears to extend beyond the ceiling.

Once the ceiling is fully retracted, the side walls of the classroom can be lowered to transform the classroom into a 360° theater (Fig. 1, *b*). To give thickness to the side walls we render four additional rectangles texture mapped with a brick texture. The 360° theater is implemented by the instructor selecting 360° still or video panoramas from a library. The ceiling and side wall retraction is controlled by the instructor through the user interface on their computer. The current stage of the ceiling and wall retraction is encoded with a scalar parameter by the instructor interface into the cloud and read by all headsets, thereby keeping the classroom configuration consistent for all students.

4.4 First-Person Video Stabilization

For the instructor to provide adequate assistance to a group of students, *XRXL* allows the instructor to assume the viewpoint of any student in the group. However, the first-person video feed acquired by the student eye camera cannot be used directly to provide situational awareness to the instructor, as its view direction changes substantially, frequently, and unpredictably as the student moves their head. This first-person video feed has to be stabilized.

One option for stabilizing the video feed is to acquire the geometry of the scene and to re-render it with a stable virtual camera. However, real-time depth and color acquisition is challenging. Furthermore, reprojecting the geometry to a stable viewpoint would lead to disocclusion errors when parts of the scene visible from the stable viewpoint were not visible from the eye camera viewpoint. This means that perfect stabilization does not only require color and depth from the eye camera viewpoint, but also from nearby viewpoints to complete the reconstruction from the stable viewpoint.

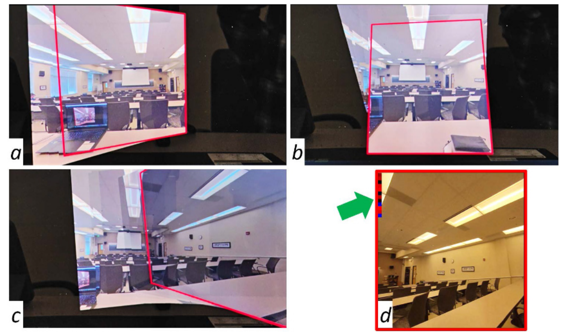

Figure 4: First-person video feed stabilization: three frames of the stabilized video (*a-c*), recorded with a phone camera aimed at the instructor laptop screen, and student headset frame (*d*, used in *c*). The current student frame is shown with the red rectangle. Although the student changes view direction considerably, the stabilized visualization shows the classroom in a consistent orientation.

A simpler alternative, which we adopt, is based on the observation that the source of instability in the first-person student video is primarily due to view direction rotations and not to viewpoint translations. Therefore, the viewpoint translation can be ignored, which allows reprojecting the eye camera frame to the stable view frame through a plane at infinity 3D homography. The homography provides a bijective mapping between the two frames, so all stabilized frame pixels can be looked up in the original frame, undoing the view rotation. The method has been used successfully in prior art, for example in the context of surgical telementoring [38].

Our video stabilization approach takes as input the intrinsic parameters of the camera and a stream of frames with known camera pose (i.e., camera extrinsic parameters). The stabilized visualization shows a dynamic trail of the k (e.g., 10, 15) most recent frames each rendered as a 3D rectangle oriented according to the frame‘s camera pose and texture mapped with the frame color (Fig. 4). Visibility is resolved with a painter‘s style algorithm with the most recent frame always winning over older frames.

We implement the streaming of the student video using Meta developer hub casting service [50], as developers are not given direct access to the passthrough video feed for privacy reasons. The stabilization runs on the instructor laptop. The implementation of the stabilization approach has to overcome several complications.

Frame and pose stream synchronization. One complication is that the frames are obtained indirectly, through casting, which incurs an unknown and variable delay. As such, it is difficult to synchronize the camera pose stream with the frame stream. We avoid the synchronization issue by embedding the quaternion defining the camera pose into the pixel data of the frame. For robustness with compression artifacts, we use three 34×34 pixel squares for each of the four floating point numbers of the quaternion. Each square carries one decimal, so we encode the quaternion with a precision of three decimals. The squares are placed at the edge of the frame and are small compared to the entire frame, see green arrow in Fig. 4. The instructor stabilization decodes the quaternion from the squares to recover the camera pose for each frame, and the squares are cropped off to not be visible in the stabilized visualization. Since the frame pose is written into the color data of the frame, the frame and pose streams are effectively glued together, without the possibility of becoming unsynchronized.

Casting camera intrinsic parameters. A second complication is that the casting crops the camera frame by undocumented amounts. The remaining field of view is not documented, and also changing frequently with updates to Meta‘s casting service. We have reverse engineered the horizontal field of view of the cropped frame through an optimization that converges on the value (i.e., 72.5°), providing a good alignment of overlapping frames.

Authorized licensed use limited to: Purdue University. Downloaded on April 26,2025 at 15:13:31 UTC from IEEE Xplore. Restrictions apply.

5 User Study

We have conducted an IRB approved user study (N = 82) to investigate our two research questions:. The participants served as students in a mock-lecture on the architecture and function of neural networks during which they were exposed to the features of *XRXL*. Specifically, to answer our **RQ1** and **RQ2** research questions (see Sec. 1), we investigated five research hypotheses.

RH1: *XRXL* can function in a large lecture with 50+ students.

RH2: *XRXL* will be judged favorably by its users.

RH3: *XRXL* is usable.

RH4: *XRXL* requires little effort from its users.

RH5: *XRXL* does not make its users cybersick.

Participants. We recruited N = 82 students from our campus with majors in which students learn about neural network architecture and function. 90% were undergraduate, and 10% were graduate students with an average age of 19.98 years. 52% were computer science majors, 20% computer technology, 18% computer engineering, and 10% data science majors. 20% of the participants were female, 80% male. Regrettably, this gender imbalance echoes the gender imbalance of the students with technical majors at our university, so our participant pool is representative of the students likely to use *XRXL* in a neural network class. 49% of the participants declared themselves familiar with neural networks, and 41% with neural network handwritten digit recognition.

Regarding familiarity with virtual reality (VR) headset technology, 12% of participants indicated that they had never used one, 22% once, 35% between two and five times, 17% more than five times, and 13% frequently. 49% of participants indicated that they had never used an XR headset (i.e., an augmented reality, mixed reality, or extended reality headset) before, 16% once, 27% between two and five times, 2% more than five times, and 6% frequently. This is expected since XR headsets at an attractive performance/price point have been available for less than a year.

Procedure. We scheduled the mock-lecture on a Saturday, for it to fit in the schedule of as many of our students as possible. We used a large classroom on our campus that can seat up to 100 students, with fixed desks, and with movable chairs (Fig. 1). The research team set up *XRXL* before the lecture, placing the XR headsets on the desks and plugging them into the desk power outlets. We did not distribute handheld controllers to simplify the logistics of having to keep track of so many similar parts being in close proximity of one another. Participants were seated, they filled out the consent form, they put on the headset, they participated in the lecture using *XRXL*, and finally they removed the headset and filled out questionnaires administered electronically through their phone. The total participant involvement took one hour, and participants were compensated with a gift card of a value equivalent to 30 USD.

The lecture lasted 30 min. One of the members of our team served as the instructor. The instructor controlled the lecture using their laptop, from the classroom lectern (Fig. 1, *d*). The lecture covered a neural network trained for hand-written digit recognition (Fig. 1, *a*). 3D visualizations showed the neural network architecture, with its layers, neurons, and connections. The instructor explained how the input image propagated through the neural network that classifies the digit correctly. The instructor pointed at the 3D visualization, scaled, rotated, and translated it, and also virtually retracted the classroom roof to accommodate a large scale version visualization of the neural network. The instructor virtually partitioned the classroom into student groups of four, each with their own neural network visualization and their own 2D panel where they answered multiple choice questions about the neural network. The participants did not have handheld controllers and were instead shown how to provide the answer with hand gestures.

The focus of this initial study was to showcase to participants the capabilities of *XRXL*, so the demonstration was not limited to neural networks. The instructor virtually retracted the walls to turn the classroom into a 360° theater that accommodated a large-scale visualization of the neural network, and that also allowed the instructor to virtually relocate the class to locations that included Times Square, under the Eiffel Tower, and an African safari (Fig. 1, *b*).

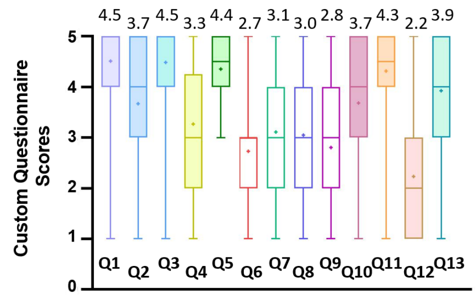

Figure 5: Box-plots of the participant five-point Likert scale answers to the custom questionnaire about experiencing *XRXL* in the large lecture. The answers of negatives questions are flipped (hollow bars, i.e., Q6–Q9). The average score is also provided numerically.

Data collection and analysis. The support for research hypothesis **RH1** is quantified objectively as the number of students for whom *XRXL* functions throughout the lecture. We investigated **RH2** through a custom questionnaire with 13 questions covering the main elements of *XRXL*, as well as the participant's speculation on whether they would want to use XR technologies in large lectures in the future. Participants provided answers on a five-point Likert scale, i.e., "strongly disagree", "disagree", "neither agree nor disagree", "agree", "strongly agree", which were analyzed using descriptive statistics. A score a of a negative question was flipped to $6-a$, for more to always mean better. The questions are : **Q1:** I enjoyed seeing the 3D visualizations. **Q2:** I prefer 3D visualizations to traditional 2D visualizations projected on the screen in front of the classroom. **Q3:** I enjoyed the classroom turning into a 360° movie theater. **Q4:** I found it easy to communicate with the other three members of my group when discussing questions. **Q5:** I like that a group can ask the instructor for help. **Q6:** I found wearing the headset to be uncomfortable. **Q7:** I did not like that we could not see each other's faces. **Q8:** I found it hard to enter the answer to the multiple choice questions using hand gestures. **Q9:** Immersive visualization is fun but it is distracting and it could get in the way of learning. **Q10:** I would like to use immersive visualization in my large lectures. **Q11:** Immersive visualization could make large lectures less boring and more engaging. **Q12:** I would consider investing $500 in a headset to bring to class. **Q13:** I would only use immersive visualization if the headset is provided to me.

The questions were selected to gauge the student's subjective perception of potential strengths and weaknesses of *XRXL*. Specifically, the questions do not avoid any potential shortcomings of the system, drawing the students' attention and requiring their input regarding the physical encumbrance of the headset, the fact that the headset hides student faces, the unfamiliar interface for entering answers through gestures, the distraction brought by XR technology, and the cost of the headset that they might be asked to cover.

RH3, **RH4**, and **RH5** were investigated using standard usability (SUS [12, 34]), task load (NASA TLX [26]), and simulator sickness (SSQ [32]) questionnaires, respectively, which are well accepted and widely used instruments. The answers were converted to a numerical score and interpreted based on the standard scales developed for each questionnaire.

Results and discussion. **RH1.** *XRXL* functioned throughout the lecture for 80 of the 82 participants, i.e., for 98% of the participants, which shows that for our user study **RH1** was supported. For one participant the headset shut down during the lecture, and for another participant the registration failed and they did not see the 3D visualization.

Authorized licensed use limited to: Purdue University. Downloaded on April 26,2025 at 15:13:31 UTC from IEEE Xplore. Restrictions apply.

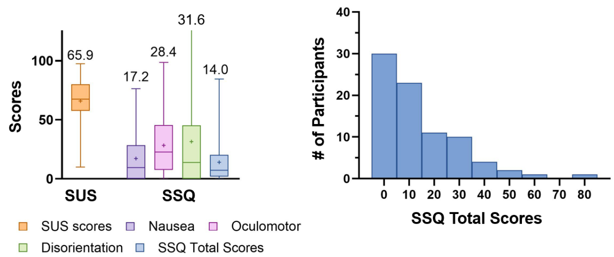

Figure 6: Left: box-plots and average participant answers to the SUS and SSQ questionnaires. Right: histogram of SSQ scores.

RH2. The answers to our custom questionnaire are given in Fig. 5. **Q1–Q3** pertain to the instructor mode. There was a strong favorable response to **Q1**, i.e., participants enjoyed seeing the 3D visualizations (4.5/5.0). The answers to **Q2** (3.7/5.0) indicate that participants have a mild preference for 3D visualizations over traditional 2D visualizations. For our mock lecture the 3D visualizations completely replaced the traditional 2D visualizations, as we wanted to give participants as much exposure to 3D visualizations as possible. It is the case that 3D visualization should not completely replace 2D visualization, but rather the two types of visualization should be interleaved throughout the lecture, for the most suitable type to be used at all times. **Q3** scores show that participants were overwhelmingly enthusiastic (4.5/5.0) about the ability of *XRXL* to turn the classroom into a 360° theater.

Q4 and **Q5** pertain to the student group mode. Participants were neutral regarding the ease of communication with their group peers, i.e., average 3.3/5.0 and median 3.0/5.0 for **Q4**. One of the obvious challenges is that since students wear headsets, they cannot see each other‘s faces. However, the average answer to **Q7** that elaborates on this point is 3.1/5.0, so students are neutral about not being able to see each other‘s faces. Unfortunately, our free Agora account quota was exceeded during the study and we could not collect meaningful responses to **Q5** during the main study. To remedy this we ran a small subsequent study with N = 14 students, which provided the strong positive answers to **Q5**, i.e., 4.4/5.0, shown in Fig. 5. Even though the form factor of XR headsets has improved, wearing them for a long time still causes some discomfort, with sub-neutral answers to **Q6** (2.7/5.0). As most participants are not used to wearing a VR headset, it is reasonable to expect that they might get more used it should they wear it routinely in class. **Q8** shows that participants did not find it exceedingly hard, nor exceedingly easy (3.0/5.0), to enter the answers to the multiple choice question on the 2D panel using hand gestures. For first time headset users, the handheld controllers would likely have had less of learning curve. We foresee that over multiple lectures the students‘ familiarity with the hand gesture interface increases, which will ultimately validate the design choice of not relying on hand-held controllers.

Q9–Q13 ask the students to speculate about their desire to use *XRXL* in actual large lectures. Based on **Q9** (2.8/5.0), students were neutral about whether immersive visualization could be a distraction that impedes learning. This opinion may be due to the students using the headsets during the whole study. Our planned use of XRXL in actual lecture entails using a mix of immersive and conventional non-immersive visualization; therefore, we believe the distraction brought by 3D visualization will be minimized while maintaining benefit of increased engagement. The openness to the inclusion of 3D visualization in large lectures is supported by the answers to **Q10** (3.7/5.0). The students were of the opinion that *XRXL* can increase engagement in large lectures and make them less boring, as shown by the 4.3/5.0 score for **Q11**. Participants are reticent to invest in a state-of-the-art XR headset, i.e., $500 for a Quest 3 [53] at the time of writing (**Q12**, 2.2/5.0). This indicates that XR headsets will have to come down in price some more before mass deployment through the preferred model where students own their headset and bring it to class. For now, immersive visualization in large lectures would have to be supported by providing the headset to students, as confirmed by **Q13** (3.9/5.0).

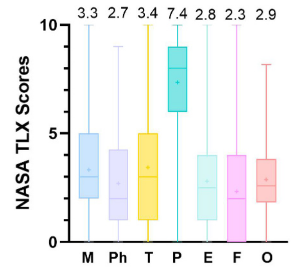

Figure 7: Box-plots of the raw NASA TLX sub-scales, and overall score: M–mental demand, Ph–physical demand, T–time demand, P–performance, E–effort, F–frustration, O–overall. P is the only positive sub-scale, i.e., more is better. For the other sub-scales less is better.

Overall, students had a favorable opinion of *XRXL*, which supports **RH2**: they greatly enjoyed seeing 3D visualizations, having the classroom turn into a 360° theater, and having the instructor provide help to individual groups. Students have also expressed some concerns. One pertains to making sure to only replace conventional 2D visualization with 3D visualization when needed. This decision can be left to the instructor, who might decide to provide an initial 3D visualization of a concept followed by a more detailed analysis on a 2D visualization. Future extensions of *XRXL* should facilitate the transition between 2D and 3D visualization, morphing the 3D visualization from and into the 2D visualization, providing continuity between 2D and 3D visualization, to avoid cognitively taxing abrupt visualization context switches. The long term concern of the encumbrance brought by wearing the headset remains. The palliative solution of limiting the headset use to short periods of time will ultimately have to be supplanted by headset form factors that approach those of vision glasses, direction in which steady progress is made (e.g., Project Aria [51], Project Orion [52]). Another concern is the headset cost, which is now $300, so the students‘ willingness to invest in a headset should be reassessed.

RH3. Fig. 6, left, shows the SUS scores. The average score of 65.9 is slightly below the 68 overall SUS average, and it maps between the adjectives “Good” and “OK” [6]. We conclude that **RH3** is partially supported. The main usability issue was the fact that the students in a group could not see the instructor video when they called for help, due to the Agora free quota being depleted. Indeed, the small subsequent study that restored the instructor video streaming ability of *XRXL* resulted in an SUS average score of 72.5, which is comfortably above the average of 68, i.e., between the 65th to 69th percentiles , and between the “Good” and “Excellent”.

RH4. Fig. 7 shows the raw NASA TLX scores for each of the 6 sub-scales, as well as overall, on a scale from 0 to 10. The averages are also shown numerically. The only positive sub-scale is P, where more performance is better, while the others measure aspect of “load” and less is better. P was inverted when computing the overall values. Participants did not find that using *XRXL* implied a high mental demand (M). Indeed, by design, *XRXL* relies on a simple interface on the student side, relying on the instructor to control the 3D visualizations. The low physical demand score (Ph) indicates that the headset discomfort reported in **Q6** is an annoyance rather than a physical burden. Participants did not feel rushed (T), they felt like they were successful in accomplishing what they were asked to do (P), that they did not need to work hard to accomplish what they were asked to do (E), and that they did so with little frustration (F). The overall task load scores were low, with an average of 2.9/10.0 (O), and we conclude that the **RH4** is supported.

RH5. Participants were instructed to remove the headset and

Authorized licensed use limited to: Purdue University. Downloaded on April 26,2025 at 15:13:31 UTC from IEEE Xplore. Restrictions apply.

to stop participating in the study at the onset of any cybersickness symptom. They were also informed that they will be compensated even if they have to stop early. No participant complained of cybersickness symptoms and stopped early. Fig. 6 shows the subscale and total SSQ scores, in numerical, boxplot, and histogram forms. The total score was computed with the original formula [32]. The average of 14.0 is low for immersive visualization applications [9], which we attribute to the high and constant frame rate of the passthrough background that anchors the users. We conclude that cybersickness is not a concern, i.e., **RH5** is supported.

6 Conclusions. Limitations. Future Work

We have presented the design, implementation, and initial validation of *XRXL*, a system for deploying immersive visualization in large lectures. Each student wears an XR headset to see 3D visualizations integrated into their view of the classroom. The instructor controls *XRXL* from their laptop, and can virtually retract the ceiling and walls of the classroom to accommodate large scale visualizations that go beyond the physical boundaries of the classroom or to turn the classroom into a 360° theater. The instructor can also virtually partition the classroom into small student groups, and the instructor can pay virtual visits to individual groups to provide assistance. We have tested *XRXL* in a mock-lecture with 82 students with promising robustness, subjective student preference, usability, task load, and cybersickness results.

Our first prototype has several limitations that future work should address. One is to simplify the setup before lecture by saving the classroom proxy for reuse. Another is to make headset to classroom alignment more robust by using multiple spatial anchors throughout the classroom, or by using the classroom proxy as opposed to the spatial anchor defined at the front of the classroom. When a student turns on the headset and specifies the classroom they are in, the headset should align automatically to the classroom. Presently, the headset manufacturer does not provide to developers direct access to the passthrough video stream, which leads to cumbersome, approximate solutions for making do with the cropped and delayed casting stream from the cloud, as detailed in Sec. 4.4. A change in this policy will be beneficial to many research and development projects like ours. With the current prototype, students do not see each other‘s faces, and future work should integrate headset inpainting solutions for virtual headset removal.

Our study investigated the unique features brought by *XRXL*, such as the large-scale visualization or the partitioning of the classroom into small groups of students. *XRXL* supports partitioning the classroom with any granularity, including with the finest possible granularity of individual students, each with their visualization, which can prove to be the visualization format of choice for some parts of the lecture and some visualization payloads. Similarly, *XRXL* is compatible with any prior art XR interaction–our study tested 0-shot hand gesture selection of multiple choice answers to be able to get rid of the daunting logistics of hundreds of handheld controllers and their batteries.

We do not advocate that students use immersive visualization exclusively, wearing the headset throughout the lecture. To support interleaving immersive and non-immersive visualization, *XRXL* has to allow for a gradual and continuous transition between a 3D visualization, seen by the student in the headset, and a 2D visualization of the same dataset, seen by the student on their laptop or on the projection screen at the front of the classroom. For example, the student should be able to put on the headset to see a visualization projected on the classroom screen come to life in 3D and then fly back to the screen, without disorienting visual discontinuities.

Our study had participants in the student role, and future studies should also investigate the instructor role. Our study inherited the gender imbalance of Computer Science courses at our institution, and future studies should deploy *XRXL* in other courses with more balanced demographics. Our initial study only began to answer feasibility, usability, and acceptability questions, and stops well short of gauging any learning benefits it might bring in the context of large lectures. Learning outcomes will have to be investigated in future longitudinal studies, run over entire semesters, with actual students and actual instructors, in actual lectures and actual courses. Controlled study designs will be needed to quantify any *XRXL* benefit in relation to conventional lectures or to other technological interventions. *XRXL* has the advantage of detailed and high frequency telemetry on each of its many students, and future work should investigate mining this data to quantify, monitor, and sustain student engagement. For example, students who do not watch the classroom or their group‘s visualization for an extended period of time are likely to be disengaged. Any such real-time data collection should be done with full consideration of privacy concerns, based on solutions and protocols yet to be developed. Students who attempt the same user interface or learning task repeatedly should be provided help in pre-recorded or automatically generated form, and, should this fail, the instructor should be notified.

An important challenge for all XR applications is content creation. In our case, *XRXL* cannot be deployed in a lecture if there is no XR ready content for the lecture. The successes of generative AI at creating text, image, video, and 3D content should be extrapolated to XR, in a way that takes into account not just shape and appearance, but also interaction and other XR specific constraints.

Headsets are still expensive, but so is a physically reconfigurable classroom, and even more so having tens of students disengage in the many large lectures on our campuses. We foresee that there are two models for campus-wide deployment of immersive visualization in large lectures. One model, more plausible for the near future, is for the institution of higher education to invest in initial deployments in a few classrooms by purchasing headsets, and by establishing specialized information technology (IT) support teams that support content creation and classroom deployment. *We note that such support is likely to be less costly and more scalable than the support for physically reconfigurable classrooms, with their expensive and fragile pivoting furniture and retractable displays.* The students walk into the classroom to find the headsets on their desks ready to go, much the same way they did in our study. This first model has the advantage of bootstrapping the deployment of immersive visualization in large lectures, and the disadvantages of poor scalability with the number of classrooms and of the lack of student buy-in. The second model is for each student to own their headset, bringing it to class, which has the advantages of students being more invested in maintaining and familiarizing themselves with their headset, and of scalability with the number of classrooms, and the disadvantage of the financial burden on students. We foresee that, ushering in this more scalable second model of student managed headsets, can be aided not just by headset cost decreases, but also by the development of online XR lesson libraries as easily accessible as online videos. Finally, the same way computer games have enormously benefited domains beyond entertainment, such as visualization in science and education, and such as the advance of AI-specific computing hardware, we foresee that the deployment of XR in large lectures will be indirectly but decisively boosted by any XR killer app, which will motivate students to overcome the headset cost and the headset interface learning curve barriers.

We are optimistic that XR headsets will soon become an important computing tool, alongside phones, laptops, and desktops, providing eloquent 3D visualizations unconstrained by the availability and size of physical displays, in education and beyond.

Acknowledgments

This material is based upon work supported by the National Science Foundation under Grants No. 2309564, 2212200, and 2417510, and by Purdue University through a donation from Eli Lilly.

Authorized licensed use limited to: Purdue University. Downloaded on April 26,2025 at 15:13:31 UTC from IEEE Xplore. Restrictions apply.

REFERENCES

[1] Photon Engine, Photon Unity Networking. https://www.photonengine.com/pun. Accessed: 2024-9-6. 5

[2] M. Adcock, S. Anderson, and B. Thomas. Remotefusion: real time depth camera fusion for remote collaboration on physical tasks. In *Proceedings of the 12th ACM SIGGRAPH international conference on virtual-reality continuum and its applications in industry*, pp. 235–242, 2013. 2

[3] Agora. Agora Real-Time Voice and Video Engagement. https://www.agora.io/en/. Accessed: 2024-9-6. 5

[4] Apple. Apple Vision Pro. https://www.apple.com/apple-vision-pro/. Accessed: 2024-9-6. 3, 4

[5] J. J. Appleton, S. L. Christenson, D. Kim, and A. L. Reschly. Measuring cognitive and psychological engagement: Validation of the student engagement instrument. *Journal of school psychology*, 44(5):427–445, 2006. 2

[6] A. Bangor, P. Kortum, and J. Miller. Determining what individual sus scores mean: Adding an adjective rating scale. *Journal of usability studies*, 4(3):114–123, 2009. 8

[7] I. Barakonyi, T. Fahmy, and D. Schmalstieg. Remote collaboration using augmented reality videoconferencing. In *Graphics Interface*, vol. 2004, pp. 89–96, 2004. 2

[8] T. Baranowski and E. J. Lyons. Scoping review of pokemon go: comprehensive assessment of augmented reality for physical activity change. *Games for health journal*, 9(2):71–84, 2020. 3

[9] P. Bimberg, T. Weissker, and A. Kulik. On the usage of the simulator sickness questionnaire for virtual reality research. In *2020 IEEE conference on virtual reality and 3D user interfaces abstracts and workshops (VRW)*, pp. 464–467. IEEE, 2020. 9

[10] M. Boekaerts. Engagement as an inherent aspect of the learning process. *Learning and instruction*, 43:76–83, 2016. 2

[11] A. Bradley, J. Klivington, J. Triscari, and R. van der Merwe. Cinematic-l1 video stabilization with a log-homography model. In *Proceedings of the IEEE/CVF winter conference on applications of computer vision*, pp. 1041–1049, 2021. 2

[12] J. Brooke et al. Sus-a quick and dirty usability scale. *Usability evaluation in industry*, 189(194):4–7, 1996. 7

[13] N. Brooks and S. Goldin-Meadow. Moving to learn: How guiding the hands can set the stage for learning. *Cognitive Science*, 40(7):1831–1849, 2016. 2

[14] B. Brucker, A.-C. Ehlis, F. B. Häußinger, A. J. Fallgatter, and P. Gerjets. Watching corresponding gestures facilitates learning with animations by activating human mirror-neurons: An fnirs study. *Learning and Instruction*, 36:27–37, 2015. 2

[15] S. Bueno, M. D. Gallego, and J. Noyes. Uses and gratifications on augmented reality games: An examination of pokémon go. *Applied Sciences*, 10(5):1644, 2020. 3

[16] Y. Cao, X. Qian, T. Wang, R. Lee, K. Huo, and 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*, pp. 1–13, 2020. 3

[17] F.-K. Chiang, X. Shang, and L. Qiao. Augmented reality in vocational training: A systematic review of research and applications. *Computers in Human Behavior*, 129:107125, 2022. 3

[18] M. R. Cowart. *A theoretical framework for evaluating embodied cognition*. The University of Wisconsin-Madison, 2000. 2

[19] S. Doukianou, D. Daylamani-Zad, and K. O'Loingsigh. Implementing an augmented reality and animated infographics application for presentations: effect on audience engagement and efficacy of communication. *Multimedia Tools and Applications*, 80(20):30969–30991, 2021. 3

[20] A. Estudante and N. Dietrich. Using augmented reality to stimulate students and diffuse escape game activities to larger audiences. *Journal of Chemical Education*, 97(5):1368–1374, 2020. 3

[21] C. J. Finelli and J. E. Froyd. Improving student learning in undergraduate engineering education by improving teaching and assessment. *Advances in Engineering Education*, 2019. 2

[22] C. Frueh, A. Sud, and V. Kwatra. Headset removal for virtual and mixed reality. In *ACM SIGGRAPH 2017 Talks*, pp. 1–2. 2017. 3

[23] R. Garrett. Inspiring engagement through creative and embodied learning. *Curriculum Studies in Health and Physical Education*, 13(3):270–283, 2022. 2

[24] M. Gattullo, A. Evangelista, A. E. Uva, M. Fiorentino, and J. L. Gabbard. What, how, and why are visual assets used in industrial augmented reality? a systematic review and classification in maintenance, assembly, and training (from 1997 to 2019). *IEEE transactions on visualization and computer graphics*, 28(2):1443–1456, 2020. 3

[25] F. Ghorbani Lohesara, K. Egiazarian, and S. Knorr. Expression-aware video inpainting for hmd removal in xr applications. In *Proceedings of the 20th ACM SIGGRAPH European Conference on Visual Media Production*, pp. 1–9, 2023. 3

[26] S. Hart. Development of nasa-tlx (task load index): Results of empirical and theoretical research. *Human mental workload/Elsevier*, 1988. 7

[27] L. C. Hodges. Student engagement in active learning classes. *Active learning in college science: The case for evidence-based practice*, pp. 27–41, 2020. 2

[28] V. Holec and R. Marynowski. Does it matter where you teach? insights from a quasi-experimental study of student engagement in an active learning classroom. 2020. 2

[29] S. Huang and V. Popescu. Hyperxrc: Hybrid in-person+ remote extended reality classroom-a design study. In *2024 IEEE Conference Virtual Reality and 3D User Interfaces (VR)*, pp. 609–618. IEEE, 2024. 3

[30] G. Iaria and H. Hubball. Assessing student engagement in small and large classes. *Transformative Dialogues: Teaching and Learning Journal*, 2(1), 2008. 2

[31] M. C. Johnson-Glenberg, C. Megowan-Romanowicz, D. A. Birchfield, and C. Savio-Ramos. Effects of embodied learning and digital platform on the retention of physics content: Centripetal force. *Frontiers in psychology*, 7:1819, 2016. 2

[32] R. S. Kennedy, N. E. Lane, K. S. Berbaum, and M. G. Lilienthal. Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness. *The international journal of aviation psychology*, 3(3):203–220, 1993. 7, 9

[33] S. Kim, G. Lee, N. Sakata, and M. Billinghurst. Improving copresence with augmented visual communication cues for sharing experience through video conference. In *2014 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 83–92. IEEE, 2014. 2

[34] Kitware, Inc. *The Visualization Toolkit User's Guide*, January 2003. 7

[35] P. Kosmas and P. Zaphiris. Embodied interaction in language learning: Enhancing students' collaboration and emotional engagement. In *Human-Computer Interaction–INTERACT 2019: 17th IFIP TC 13 International Conference, Paphos, Cyprus, September 2–6, 2019, Proceedings, Part II 17*, pp. 179–196. Springer, 2019. 2

[36] G. Lampropoulos. Augmented reality and artificial intelligence in education: Toward immersive intelligent tutoring systems. In *Augmented reality and artificial intelligence: The fusion of advanced technologies*, pp. 137–146. Springer, 2023. 3

[37] J. J. Lee. Size matters: An exploratory comparison of small-and large-class university lecture introductions. *English for specific purposes*, 28(1):42–57, 2009. 2

[38] C. Lin, E. Rojas-Munoz, M. E. Cabrera, N. Sanchez-Tamayo, D. Andersen, V. Popescu, J. A. B. Noguera, B. Zarzaur, P. Murphy, K. Anderson, et al. How about the mentor? effective workspace visualization in ar telementoring. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 212–220. IEEE, 2020. 2, 6

[39] R. Lindgren, M. Tscholl, S. Wang, and E. Johnson. Enhancing learning and engagement through embodied interaction within a mixed reality simulation. *Computers & Education*, 95:174–187, 2016. 2

[40] F. Liu, M. Gleicher, H. Jin, and A. Agarwala. Content-preserving warps for 3d video stabilization. In *Seminal Graphics Papers: Pushing the Boundaries, Volume 2*, pp. 631–639. 2023. 2

[41] P. Liu, C. Li, C. Xiao, Z. Zhang, J. Ma, J. Gao, P. Shao, I. Valerio, T. M. Pawlik, C. Ding, et al. A wearable augmented reality navigation system for surgical telementoring based on microsoft hololens. *Annals of biomedical engineering*, 49:287–298, 2021. 3

[42] Y.-L. Liu, W.-S. Lai, M.-H. Yang, Y.-Y. Chuang, and J.-B. Huang.

Authorized licensed use limited to: Purdue University. Downloaded on April 26,2025 at 15:13:31 UTC from IEEE Xplore. Restrictions apply.

Hybrid neural fusion for full-frame video stabilization. In *Proceedings of the IEEE/CVF international conference on computer vision*, pp. 2299–2308, 2021. 2

[43] W. H. Lo. *ARSpectator—Enriching On-Site Sports Spectating with Augmented Reality*. PhD thesis, University of Otago, 2022. 3

[44] W. H. Lo, S. Zollmann, and H. Regenbrecht. Xrspectator: Immersive, augmented sports spectating. In *Proceedings of the 27th ACM Symposium on Virtual Reality Software and Technology*, VRST '21. Association for Computing Machinery, New York, NY, USA, 2021. doi: 10.1145/3489849.3489930 3

[45] H. M. Marks. Student engagement in instructional activity: Patterns in the elementary, middle, and high school years. *American educational research journal*, 37(1):153–184, 2000. 2

[46] M. McGill, J. Gugenheimer, and E. Freeman. A quest for co-located mixed reality: Aligning and assessing slam tracking for same-space multi-user experiences. In *Proceedings of the 26th ACM Symposium on Virtual Reality Software and Technology*, pp. 1–10, 2020. 2

[47] R. R. McKnight, C. A. Pean, J. S. Buck, J. S. Hwang, J. R. Hsu, and S. N. Pierrie. Virtual reality and augmented reality—translating surgical training into surgical technique. *Current reviews in musculoskeletal medicine*, 13:663–674, 2020. 3

[48] V. M. Memmesheimer and A. Ebert. Scalable extended reality: A future research agenda. *Big Data and Cognitive Computing*, 6(1):12, 2022. 3

[49] M. Menekse, G. S. Stump, S. Krause, and M. T. Chi. Differentiated overt learning activities for effective instruction in engineering classrooms. *Journal of Engineering Education*, 102(3):346–374, 2013. 2

[50] Meta. Meta Quest Developer Hub. https://developers.meta.com/horizon/documentation/unity/ts-odh/. Accessed: 2024-9-6. 5, 6

[51] Meta. Project Aria. https://www.projectaria.com. Accessed: 2024-12-26. 8

[52] Meta. Project Orion. https://about.meta.com/realitylabs/orion. Accessed: 2024-12-26. 8

[53] Meta. Quest 3: Mixed Reality Headset. https://www.meta.com/quest/quest-3/. Accessed: 2024-9-6. 1, 3, 4, 8

[54] Meta. Quest 3S: Mixed Reality Headset. https://www.meta.com/quest/quest-3s/. Accessed: 2024-12-26. 3, 4

[55] Meta. Quest Pro: Meta Most Advanced New Virtual Reality Headset. https://www.meta.com/quest/quest-pro. Accessed: 2023-02-01. 3

[56] Meta. Scene Model API. https://developer.oculus.com/documentation/unity/unity-scene-overview/. Accessed: 2024-9-6. 5

[57] Microsoft. HoloLens 2 - Augmented Reality Headset . https://www.microsoft.com/en-us/hololens/hardware/. Accessed: 2024-9-6. 4

[58] D. Mikułowski and J. Brzostek-Pawłowska. Multi-sensual augmented reality in interactive accessible math tutoring system for flipped classroom. In *International Conference on Intelligent Tutoring Systems*, pp. 1–10. Springer, 2020. 3

[59] P. D. Milanović, I. V. Popadić, and B. D. Kovačević. Gyroscope-based video stabilization for electro-optical long-range surveillance systems. *Sensors*, 21(18):6219, 2021. 2

[60] K. C. Millett. Making large lectures effective: An effort to increase student success. In *The teaching and learning of mathematics at university level: an ICMI study*, pp. 137–152. Springer, 2001. 2

[61] A. N. Nagele, V. Bauer, P. G. Healey, J. D. Reiss, H. Cooke, T. Cowlishaw, C. Baume, and C. Pike. Interactive audio augmented reality in participatory performance. *Frontiers in Virtual Reality*, 1:610320, 2021. 3

[62] V. Pamuru, W. Khern-am nuai, and K. Kannan. The impact of an augmented-reality game on local businesses: A study of pokémon go on restaurants. *Information Systems Research*, 32(3):950–966, 2021. 3

[63] D. Parmar and T. Bickmore. Making it personal: Addressing individual audience members in oral presentations using augmented reality. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 4(2):1–22, 2020. 3

[64] R. Raj, P. Rajiv, P. Kumar, M. Khari, E. Verdú, R. G. Crespo, and G. Manogaran. Feature based video stabilization based on boosted haar cascade and representative point matching algorithm. *Image and Vision Computing*, 101:103957, 2020. 2

[65] M. Ramchander and M. J. Naude. The relationship between increasing enrolment and student academic achievement in higher education. *Africa Education Review*, 15(4):135–151, 2018. 2

[66] H. Regenbrecht and T. Schubert. Measuring presence in augmented reality environments: design and a first test of a questionnaire. *arXiv preprint arXiv:2103.02831*, 2021. 2

[67] E. Rojas-Muñoz, C. Lin, N. Sanchez-Tamayo, M. E. Cabrera, D. Andersen, V. Popescu, J. A. Barragan, B. Zarzaur, P. Murphy, K. Anderson, T. Douglas, C. Griffis, J. McKee, A. Kirkpatrick, and J. P. Wachs. Evaluation of an augmented reality platform for austere surgical telementoring: a randomized controlled crossover study in cricothyroidotomies. *NPJ digital medicine*, 3(1):75, 2020. 3

[68] A. Scavarelli, A. Arya, and R. J. Teather. Virtual reality and augmented reality in social learning spaces: a literature review. *Virtual Reality*, 25:257–277, 2021. 3

[69] A. Schäfer, G. Reis, and D. Stricker. A survey on synchronous augmented, virtual, andmixed reality remote collaboration systems. *ACM Computing Surveys*, 55(6):1–27, 2022. 2

[70] M. Sereno, X. Wang, L. Besançon, M. J. Mcguffin, and T. Isenberg. Collaborative work in augmented reality: A survey. *IEEE Transactions on Visualization and Computer Graphics*, 28(6):2530–2549, 2020. 2

[71] A. Skulmowski and G. D. Rey. Embodied learning: introducing a taxonomy based on bodily engagement and task integration. *Cognitive research: principles and implications*, 3(1):1–10, 2018. 2

[72] P. Wang, X. Bai, M. Billinghurst, S. Zhang, X. Zhang, S. Wang, W. He, Y. Yan, and H. Ji. Ar/mr remote collaboration on physical tasks: a review. *Robotics and Computer-Integrated Manufacturing*, 72:102071, 2021. 2

[73] M.-H. Wu. The applications and effects of learning english through augmented reality: A case study of pokémon go. *Computer Assisted Language Learning*, 34(5-6):778–812, 2021. 3

[74] J. Yu and R. Ramamoorthi. Learning video stabilization using optical flow. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 8159–8167, 2020. 2

[75] K. Yu, U. Eck, F. Pankratz, M. Lazarovici, D. Wilhelm, and N. Navab. Duplicated reality for co-located augmented reality collaboration. *IEEE Transactions on Visualization and Computer Graphics*, 28(5):2190–2200, 2022. doi: 10.1109/TVCG.2022.3150520 2

Authorized licensed use limited to: Purdue University. Downloaded on April 26,2025 at 15:13:31 UTC from IEEE Xplore. Restrictions apply.