Educational Telepresence Project Proposal

NICK REWKOWSKI, NICK1@UMD.EDU

ACM Reference Format:
Nick Rewkowski, nick1@umd.edu. 2020. Educational Telepresence Project Proposal. 1, 1 (October 2020), 7 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 DESCRIPTION OF THE PROJECT

1.1 Purpose

This project is meant to further understanding of how people learn in an immersive virtual environment (VE) as compared to in-classroom learning and remote learning through video conferencing, especially when immersed through a stereoscopic VR or AR HMD.

Based on prior research on spatial cognition and memory palaces (to be cited in detail later), the spatial context of an immersive VE should generally improve memory ability compared to non-spatial learning methods such as distance learning with Zoom, as the brain naturally draws connections between learned detail and the context within which is was learned with respect to the senses, personal situation, etc. Methods like Zoom learning do not have much context within which to learn, and recent post-quarantine research attempts to draw a connection between this fact and the general learning difficulties that teachers are noticing in their remote classrooms–students are having trouble focusing, retaining material, studying efficiently, etc. We suspect that VR/AR's spatial context can improve these issues and may be suitable as an alternative to video conferencing as a learning method, especially in classes in which being able to see the real faces of one's peers is not important. Prior work on education in AR and VR (such as the descriptions provided in the CHI '20 submission) also supports this concept, and recent attempts at pseudo-spatial video conferencing, like Microsoft Teams' "Together Mode" are finding success here as well.

At a high level, this requires some system that supports VR, perhaps built with a game engine, that allows students and teachers to (1) join a classroom VE with each other, (2) choose 3D avatars that they control that are displayed to others in the VE, (3) use an interface onto which class materials are presented, and (4) take notes and responds to evaluations such as quizzes.

As an example of an expected effect, a student immersed in VR can hear spatialized audio coming from people in the classroom. Spatialized audio is known to improve immersion, and the spatial memory of where a peer or the teacher were when talking provides more context that the brain can use to memorize material. The modern alternative–looking at a gray box with someone's name (or, much less likely, a face from a webcam) on Zoom and hearing someone talking–is not as memorable.

We suspect similar effects with social VR. Teachers on Zoom have had difficulty facilitating natural discussion between remote students, usually resorting to Zoom's "breakout rooms," which place a group of students into a independent sub-video conference where they cannot be heard outside of the group, usually resulting in the students not doing anything during a designated discussion time. However, recent efforts at social VR, such as AltSpace VR and VRChat, show that anonymized users in a 3D VE have no such difficulty striking up a discussion. The social VR method of holding these discussions has numerous benefits, such as spatial factors like the spatialized audio effect mentioned above providing more context for memory retention, the ability to listen to what other groups are saying (allowing for realistic cross discussion like in in-person classes), the ability for the teacher to join in on groups more easily, and so on.

The exact effect of any of these factors is not well understood, and so the major goals of this project are not only to study these factors that may/may not improve learning through VR, but to build a framework allowing for such studies to be easily held, controlled, and analyzed through automated data collection.

A more secondary goal is to develop more natural input modes for serious VR applications such as education, meetings, etc. As will be described in more detail later, there are multiple proofs of concepts of immersive interfaces for AR and VR, such as Microsoft's "Maquette"/"MRTK" and Oculus' "Infinite Office" demos, which all propose alternative input methods for interfaces such as keyboards and drawing to allow for longer, more comfortable VR work sessions. However, their proposed features have not been developed or studied in a serious context yet.

1.2 Applications

Full Semester VR Classes. Previous attempts at VR classrooms were generally limited to one or two sessions, mostly due to equipment price, uncomfortability, limited application features and context, being limited to a short-term study, etc. We hope to build something that would allow a full class to be taught in VR, especially a class that would benefit from full 3D learning environments. A 3D game development or AR/VR class is a possible application, and instructors in applications like AltSpace VR have already found success in holding such classes in VR, even successfully monetizing them in some cases due to their success. A class like this, which teaches 3D concepts such as meshing, animation, 3D audio, etc., would benefit from the teacher's ability to play with these concepts in realtime in class to improve the students' understanding (e.g. 3D modelling in realtime or turning audio propagation on/off so students immediately notice the effects). Other classes that could benefit from this are demo-heavy applications classes (such as UNC's Serious Games) or spatial math classes like Calculus 3 or Numerical Analysis. If the memory palace theories hold up, then generic non-spatial classes

Author's address: Nick Rewkowski, nick1@umd.edu.

XXXX-XXXX/2020/10-ART $15.00
https://doi.org/10.1145/nnnnnnn.nnnnnnn

may also find success here (such as the education VR attempts by Stephen King at UNC and Lenovo's K-12 VR Classroom).

Remote Learning. While COVID is a good example of why remote learning can be a necessity, there are other less common use cases, such as for students with disabilities who cannot be physically in class for whatever reason. Even for an in-person class, it is possible to track the teacher and other students with frameworks such as OpenPose that require no VR equipment for the tracked people, so that they can be reconstructed for the remote viewer to experience and interact with. Hypothetically, a teacher on an online education site like Coursera could even record their entire 3D lectures for future consumption without needing to re-record it for every session, making many difficult classes more accessible to visual learners.

More Accessible and Effective Mobile AR/VR. By improving AR/VR interaction methods and having a framework that allows for serious applications to run on mobile devices (e.g. Google Cardboard/smartphones, Oculus Quest, or HoloLens 2), we can improve the state of mobile AR/VR and make it less tedious to interact with the VE with cheaper equipment. For example, with Google Cardboards, the user types by looking at the button they want to press and hitting a button on the side of the HMD, which gets tedious very quickly, especially when trying to take notes or write full sentences. However, by utilizing the phone's rear camera for handtracking, we can make it much easier and more natural to type.

2 SPECIFIC RESEARCH QUESTIONS

In general, we want to know if the immersive effects of AR/VR improve various cognitive factors, such as memory retention, skill transfer, willingness to continue using AR/VR for learning, etc. While we do need to build a system to run such studies, this is ultimately an evaluation-based project. We only need to build the system because there are no open-source alternatives to the state-of-the-art social VR applications currently used for education, such as Altspace VR.

This section contains some research questions that we plan to study.

2.1 Basic

These are the questions that are most immediately answerable with a barebones system.

(1) Does learning in AR/VR affect memory quality and retention?
(2) How do we facilitate effective note-taking in 3-space that is as natural as taking notes in an in-class session?
(3) If we have the best note-taking features possible in VR, will we see students take more notes compared to on Zoom?
(4) How effective is it to learn from a full 3D pre-recording of an immersive lecture compared to a prerecorded Zoom lecture or Youtube tutorial?
(5) Does it matter if we simulate simple cues like blinking and lip sync on the virtual avatars of your peers? Does not having these disrupt the learning process?

2.2 Advanced

These are some more advanced questions that can be answered relatively easily if we have a modular system framework facilitating the new features and data.

(1) Can we have full hand-tracking-based note-taking with cheap mobile VR devices (e.g. with rear camera)? How effective is this compared to the state of the art mobile VR method of clicking keys with a controller/HMD button?
(2) Is it noticeable if some students have a high-end HMD that allows tracking of hands compared to students who are only using low-end devices?
(3) Does it matter if the classes are in AR vs. VR? E.g. would students prefer seeing holograms of their peers/teacher and see the real world note-taking supplies?
(4) How do we effectively replicate emotional responses in a class, e.g. in a social science or discussion-based class? This could apply to gestures, gait, and facial animations (e.g. Oculus' emotive avatars). This is especially important for classes targetting children, who are currently struggling to build social skills with remote learning.
(5) Are immersive "breakout sessions" more effective than Zoom-type breakout sessions?

3 DEFINING THE SYSTEM

3.1 Basic

These are the features we would consider to be most basic for a functioning study system that can answer the basic questions above.

(1) Ability to join a class in VR or without VR. "Without VR" may refer to joining the 3D VE as a first-person/third-person avatar or video conferencing into the class somehow. This is necessary to make any kind of statistically-valid comparison.
(2) Ability to join an ongoing class session, e.g. Kahoot-style. We can assume, for the basic questions, that there is only 1 classroom in existence that is running on a server of the teacher's or the study proctor's.
(3) Basic networking. This includes being able to move your avatar around and have it and its motions be visible to others, others being able to hear your microphone input, others being able to see what you write (if you choose to allow them to), etc.
(4) Being able to choose a seat and sit there. This is to allow relatively natural user placement as opposed to people standing around randomly like in Altspace VR.
(5) Basic avatar customization, e.g. hair/skin/eye/clothes color. Textures beyond the character facial features are unnecessary at first.
(6) A classroom VE for people to join into that has seats, presentation screen, whiteboards, etc.

3.2 Secondary

(1) Ability for the teacher and students to open personal web browsers to present on or use for research. They should be able to open/close it at will, type/click on it, resize it, and choose if others can see it. As it is possible to have discussion or seminar-style classes

in which the students should not use personal devices, this is considered secondary for now.

(2) The above feature requires some kind of keyboard/mouse simulation, such as through controller input or gaze+button input.

(3) Some method for students to take notes, e.g. drawing on surfaces or typing into something. This feature may not be necessary for the most basic questions, but is still a fundamental feature of the educational framework.

3.3 Advanced

(1) Avatars that blink and lipsync based on mic input
(2) Hand tracking using the rear phone camera, which can then be used for notes
(3) Full 3D recordings of lectures and discussions to be replayed in an immersive VE
(4) Microphone denoising; noise is currently a major problem with Zoom.
(5) Ability for guests such as parents, TAs, etc. to join
(6) Real face of the user mapped onto the virtual avatar
(7) Ability to decorate the classroom
(8) Easy way for the teacher to point at and highlight objects or people (such as when someone asks a question or a specific piece of lecture material is important).
(9) Emoji system for quick, non-mic responses
(10) Simple hand-raising ability that does not rely on the student mic input

4 SUBMISSION GOALS

The initial plan was to submit something to SIGCHI '21, but this was impossible without IRB approval, so a background literature review was submitted instead.

The next goal is the IEEE VR Conference, which is due Nov 13.

The next possible deadlines are I3D due in December (if it exists this year) or ICSC (International Conference on Spatial Cognition) due on Jan 31. It is not clear if SIGGRAPH would accept any of this, even a poster.

5 SOME INSPIRATIONAL FEATURES FROM OTHER APPLICATIONS

Descriptions in figure captions below.

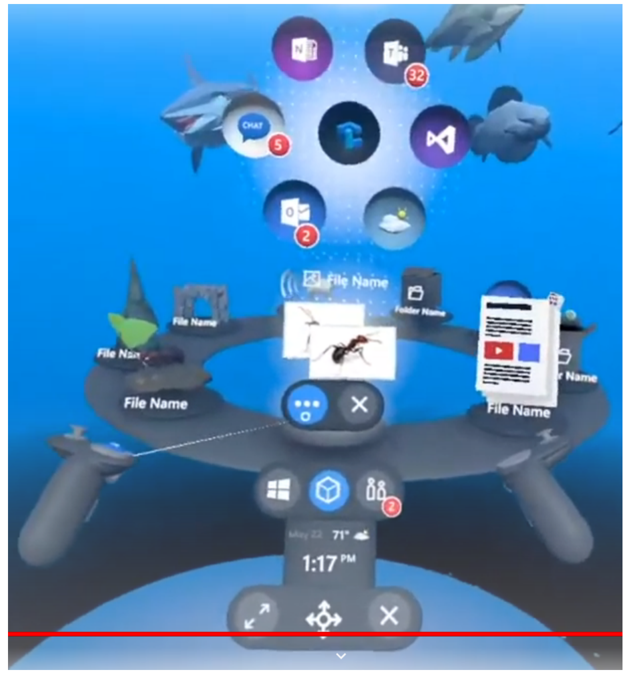

Fig. 1. Microsoft Maquette: proof-of-concept for an app selection screen

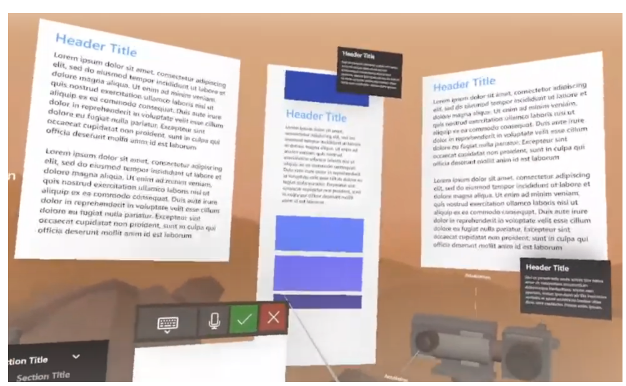

Fig. 2. Microsoft Maquette: proof-of-concept for pinned spatial documents

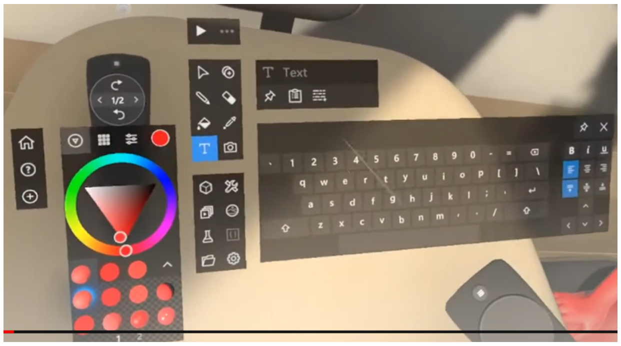

Fig. 3. Microsoft Maquette: User interface, including keyboard and color drawing methods.

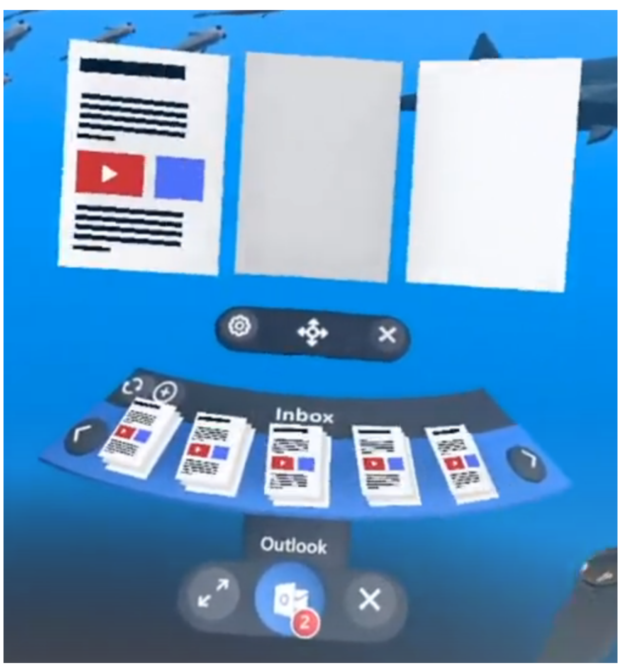

Fig. 4. Microsoft Maquette: proof-of-concept for VR email checkers

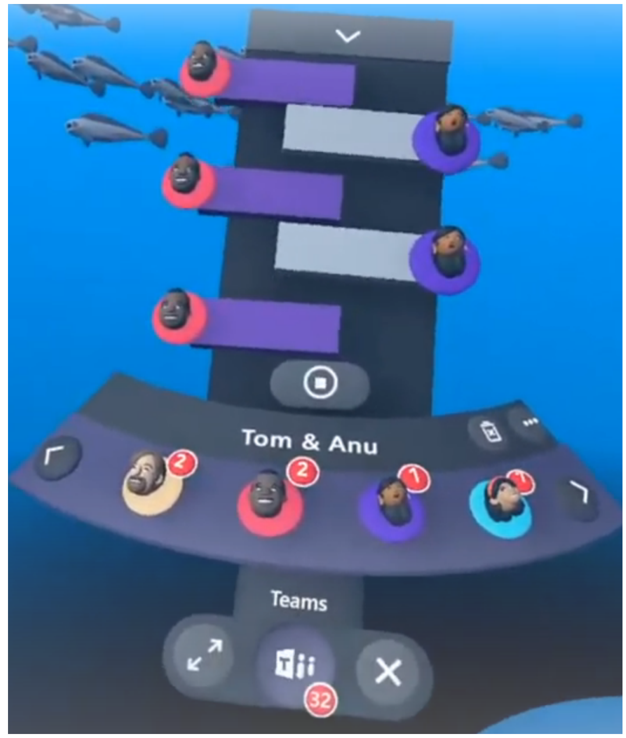

Fig. 5. Microsoft Maquette: proof-of-concept for VR messaging app

Fig. 6. Microsoft Maquette: example of spatial drawing

Fig. 7. Microsoft Maquette: proof-of-concept for a note-taking interface in VR (in particular, with OneNote)

Fig. 9. Microsoft Maquette: proof-of-concept of a remote user sitting

Fig. 8. Microsoft Maquette: proof-of-concept for a note-taking interface in VR (in particular, with OneNote) (another view)

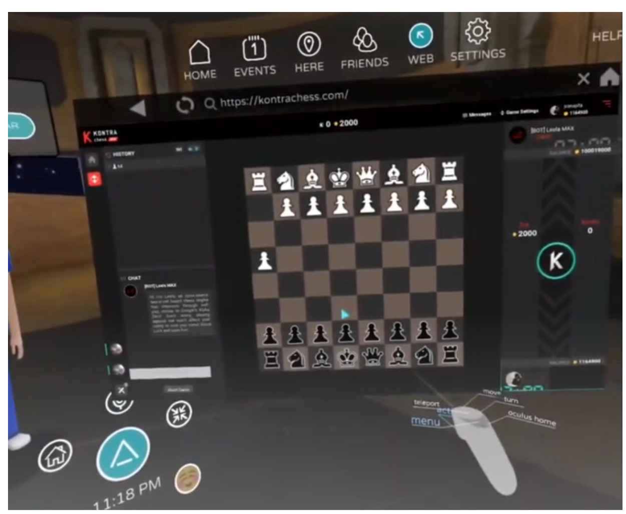

Fig. 10. AltspaceVR: openable 3D browser

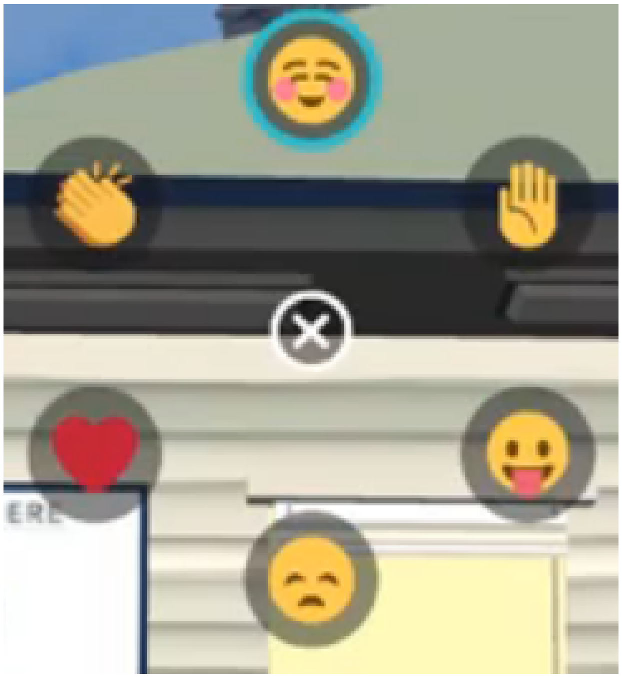

Fig. 11. AltspaceVR: emojis that you can choose from, which will then appear over your avatar's head

Fig. 12. AltspaceVR: their user control panel at the bottom left, which opens the web browser, emoji chooser, mute button, settings, home VE, avatar editor, and capture camera

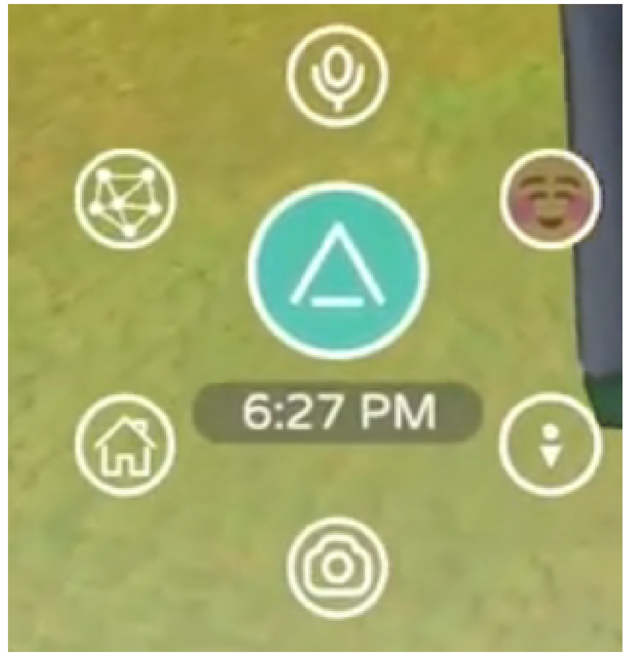

Fig. 13. AltspaceVR: closeup of the control panel

Fig. 14. AltspaceVR: example of presentation screen

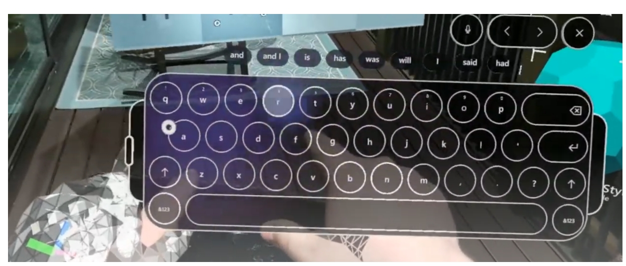

Fig. 15. HoloLens 2 Mixed Reality Toolkit (MRTK): keyboard accepting hand tracking input

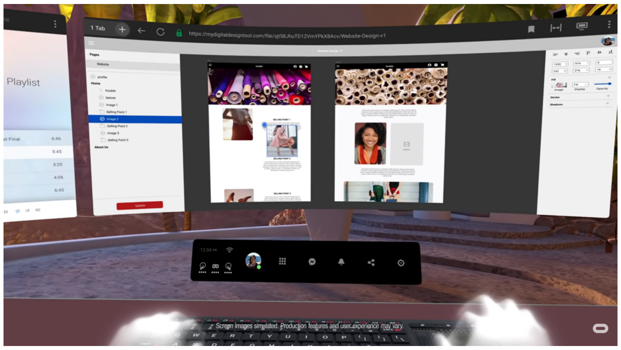

Fig. 16. Oculus' Infinite Room proof-of-concept allowing an AR view of a real keyboard for typing.