Received: 4 May 2022 | Revised: 15 May 2022 | Accepted: 15 May 2022

DOI: 10.1002/cav.2082

SPECIAL ISSUE PAPER

WILEY

Holographic sign language avatar interpreter: A user interaction study in a mixed reality classroom

Fu-Chia Yang | Christos Mousas | Nicoletta Adamo

Department of Computer Graphics Technology, Purdue University, West Lafayette, Indiana, USA

Correspondence
Fu-Chia Yang, Department of Computer Graphics Technology, Purdue University, West Lafayette, IN 47907, USA.
Email: yang1684@purdue.edu

Funding information
Purdue University, Instructional Innovation Grant

Abstract

We explored user interactions with a holographic sign language interpreter in a mixed reality (MR) classroom for deaf and hard of hearing students. The developed MR application projects a holographic signing avatar that translates in real time the lecture while a speaking instructor is teaching. Our study explored user interaction with the MR system, intending to provide design guidelines for digital MR sign language interpreters. We recruited eight participants and conducted a usability test focused on avatar framing (full-body vs. half-body) and avatar manipulation (fixed position, scale, and orientation vs. user-adjustable position, scale, and orientation) in the MR classroom. We used a mixed-method approach to analyze quantitative and qualitative data through recordings, surveys, and interviews. The results show user preferences toward viewing holographic signing avatars in the MR environment and user acceptability toward such applications.

KEYWORDS

holographic avatar, mixed reality, sign language animation, user interaction

1 | INTRODUCTION

Mainstream schools assign sign language interpreters to classes where deaf and hard of hearing (DHH) students learn with hearing students. Training a sign language interpreter for mainstream education requires years of work. Experienced interpreters are more challenging to find when it comes to higher education. Moreover, hiring trained interpreters for classes can be costly and interpreters are not easily accessible. For this reason, we explored the use of virtual sign language interpreters on a portable mixed reality (MR) device, Microsoft's HoloLens, as a promising alternative.

Since HoloLens has not been widely used compared to mobile augmented reality (AR) applications, design guidelines, and user tendencies have not been fully explored, especially for the specific user group, the DHH community. Constantinou et al.[1] suggested that more research should provide unified tools for teachers to adapt to DHH students' needs. To achieve that, assessments of DHH students' user tendencies, preferences, and human-centered interaction factors need to be conducted. Zirzow[2] and Saladin[3] stated that the literacy level of DHH individuals could be improved through access to integrating technologies. Using figurative presentations and interfaces instead of literal presentations can also strengthen DHH students' learning efficacy. By enhancing computer graphic technologies among different wearable devices, we can achieve wider diversity of computer-powered assistive technology for DHH education.

We developed a HoloLens powered MR application and conducted mixed-method usability tests with eight study participants in a recreated classroom setup. Our study sought to find the preferred MR avatar configuration for DHH students, with a holographic sign language interpreter embedded in the MR space. We implemented and tested two aspects of the holographic sign language interpreter in the MR classroom. First, sign language avatar's framing method (full-body vs. half-body). Second, the sign language avatar's manipulation settings (fixed position, orientation, and scale vs. user-adjustable position, orientation, and scale). Based on the results of this study, we provided insights and design guidelines for the future development of holographic sign language interpreters in MR classrooms.

2 RELATED WORK

Throughout the years, researchers have been working on sign language digitalization through computer animation technology. Huenerfauth and Hanson[4] discussed several computer interfaces and communication systems adopting digital sign language animations for the deaf. Kipp et al.[5] presented a gloss-based cyclic animating and evaluating approach to achieve state-of-art performance on digital sign avatars. Kaneko et al.[6] combined motion captured data to generate Japanese Sign Language animation through the TV program Making Language (TVML). Lee and Kuni[7] introduced a system translating the natural text into sign language through a hand motion coding method. Kaur and Kumar[8,9] designed a sign animation system for sign language automation using conversion between Hamburg Notation System (HamNosys) and Signing Gesture Markup Languages (SiGML). Kang[10] proposed a framework utilizing speech recognition, natural language processing, and 3D animation techniques to translate Chinese and English text into sign animation. Lastly, Vesel and Robillard[11] introduced a Signing Math Dictionary (SMD) to help deaf children learn mathematic vocabulary through digital sign animation. Their research showed enhancement in DHH students' learning efficacy through adopting digital signing avatars.

We can see from various studies that MR wearable devices have started to reveal their capability in helping DHH communities in multiple aspects. Peng et al.,[12] Jain et al.,[13] and Guo et al.[14] introduced holographic speech bubbles, an MR application that provides real-time generated speech captions that reinforce DHH people in group conversations. Vinayagamoorthy et al.[15,16] explored user preference in different positions, sizes, and scales of holographic sign language avatars for TV show interpretation. The previously mentioned research focuses on daily activities and communication purposes. However, there is also research that introduces similar technologies into DHH education. The study conducted by Adamo-Villani and Anasingaraju[17] used the holographic sign language avatar for K-6 DHH math education, powered by Meta 1 Dev Kit. Miller et al.[18] conducted a study on a similar protocol with Adamo-Villani and Anasingaraju but with Google Glasses and EPSON Moverio BT-200 glasses. Moreover, Parton[19] conducted a study using Glass Vision 3D, an AR education tool powered by Google Glasses. The previous research demonstrated the scalability and cost-effectiveness of computer-generated sign animation through wearable AR devices. The authors also discovered that DHH students showed enthusiastic engagement toward using wearable devices for learning.

The above studies provide the basis for this research. However, the limitation observed is that there's not much intention given mainly to the study of classroom arrangements for holographic sign language interpreters utilizing cutting-edge MR technology. So far, researchers have not introduced HoloLens to education for the DHH community. According to Constantinou et al.,[1] current studies on new assistive technologies for DHH remain at the prototyping stage without discovering specific requirements or suitable arrangements. Thus, our research attempts to explore the potential of deploying cutting-edge MR technologies while also investigating user acceptability toward such applications.

3 IMPLEMENTATION DETAILS

We built an MR application for DHH students to explore user interactions in an MR classroom setup. The implementation pipeline consists of three steps (see Figure 1). First, we collected four K-1 Math lessons through a motion capture session. A professional deaf signer was recruited to sign in ASL four chapters of K-1 math lectures. We utilized motion capture (see Figure 2) technology to capture ASL animation for the holographic avatar. We used the OptiTrack, StretchSense, and FaceWare software to capture the body, hands, and face, respectively. The motion data captured from each software were transferred into Autodesk's Motion Builder. Second, we post-edited the raw motion data in Autodesk Maya to fix mesh intersections and remove noise. Third, we designed the MR user interface and developed the Unity game engine application with Microsoft MRTK plugins. The Object Manipulator toolkit from MRTK was used to allow participants to adjust the sign avatar's position, orientation, and size (see Figure 3 for avatar manipulation). The air tap interaction allows participants to select objects from a distance (see Figure 4 for air tap hand gesture).

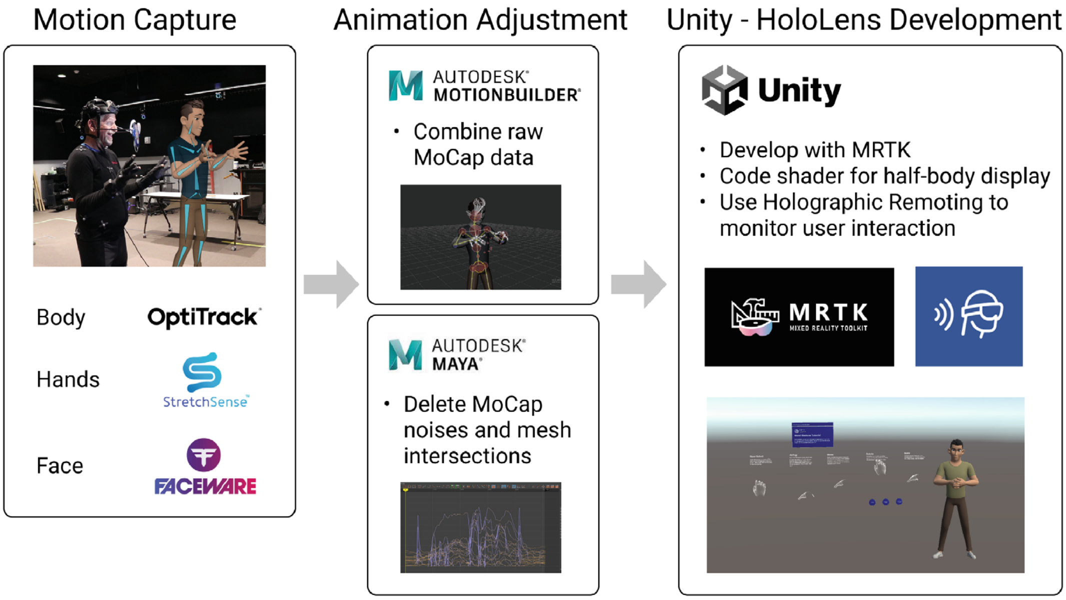


FIGURE 1 Implementation pipeline of our project

FIGURE 2 The motion capture session

4 | MATERIALS AND METHODS

4.1 | Participants

We recruited eight participants (age: $M = 22.38$, SD $= 4.18$) in our study. Seven were female and one was male. None of them were deaf or hard of hearing, two were children of deaf adults (CoDA) who were native ASL users, one used ASL for more than 10 years, one used ASL for 8 years, and four used ASL for less than 5 years. Three had experience working as ASL interpreters. Two used ASL daily, four used ASL a few times a week, and two used ASL a few times a month. Among all eight study participants, none of them had experience using MR HMDs. Three stated that they have experience using the VR HMD, Oculus.

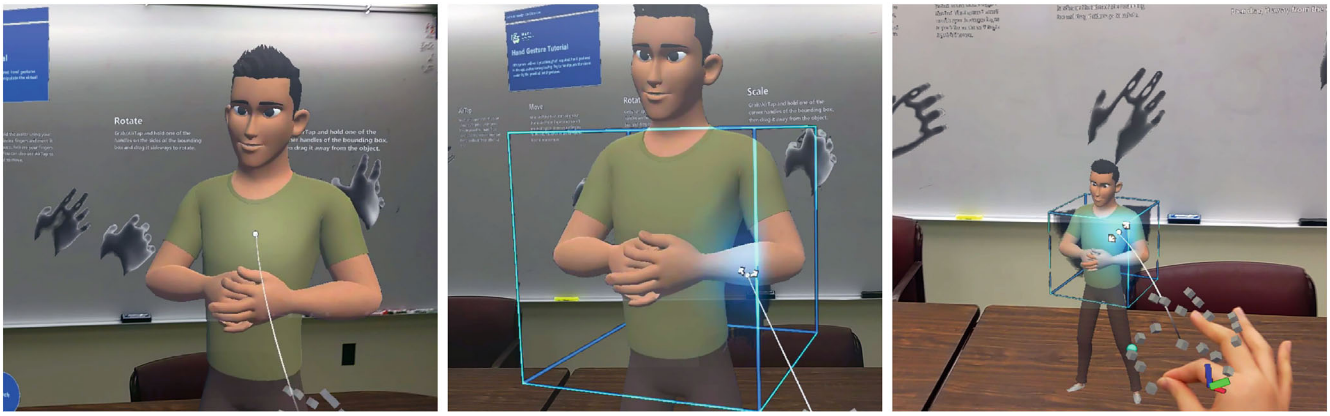


FIGURE 3 Adjusting avatar's position, orientation, and size in MR

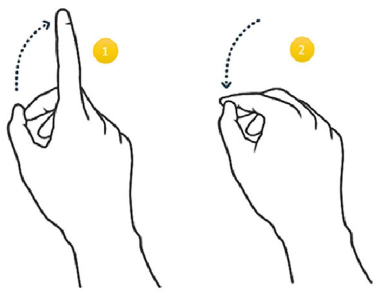


FIGURE 4 Air tap select hand gesture

4.2 | Experimental conditions

For our study, we followed a 2 (avatar framing: full-body (F1) vs. half-body (F2) display) × 2 (avatar manipulation: fixed (M1) vs. user-adjustable (M2) position, orientation, and scale) within-group study design. Thus, each participant experienced four experimental conditions (see Figure 5): M1×F1, M2×F1, M1×F2, and M2×F2.

4.3 | Measurements and ratings

This research was conducted through a mixed-method approach. Both quantitative and qualitative data were collected through testing sessions.

We collected measurements on several different aspects of user experiences and interactions. We required participants to rate the following statements on a 7-point Likert scale (1 = strongly disagreed and 7 = strongly agreed): I was very satisfied with the experience of this scenario (repeated after all four experimental conditions); (PC1) I found the HoloLens comfortable to wear; (PC2) I did not experience any motion sickness or nausea; (PC3) I could easily navigate the environment while wearing the headset; (OE1) I could do all functions I desired; (OE2) I have a high acceptance rate toward this application; (HG1) The hand gesture tutorial was helpful; (MA1) I could near select properly; (MA2) I could air tap properly; (MA3) I could move the avatar properly; (MA4) I could rotate the avatar properly; (MA5) I could resize the avatar properly. Also, they were asked provide their ratings on the NASA Task Load Index (NTLX) scale.[20]

Besides the quantitative ratings from the previously mentioned questions, the participants were asked to elaborate on their thought processes behind each rating. They were asked to share their preferred position, orientation, and scale of the avatar in the MR space; their preferences between the proposed four experimental conditions. The open-ended interview asked whether they think this application would be helpful to DHH students, what are the preferred additional features, and what other situations can benefit from this application. We used content and narrative analysis to code data

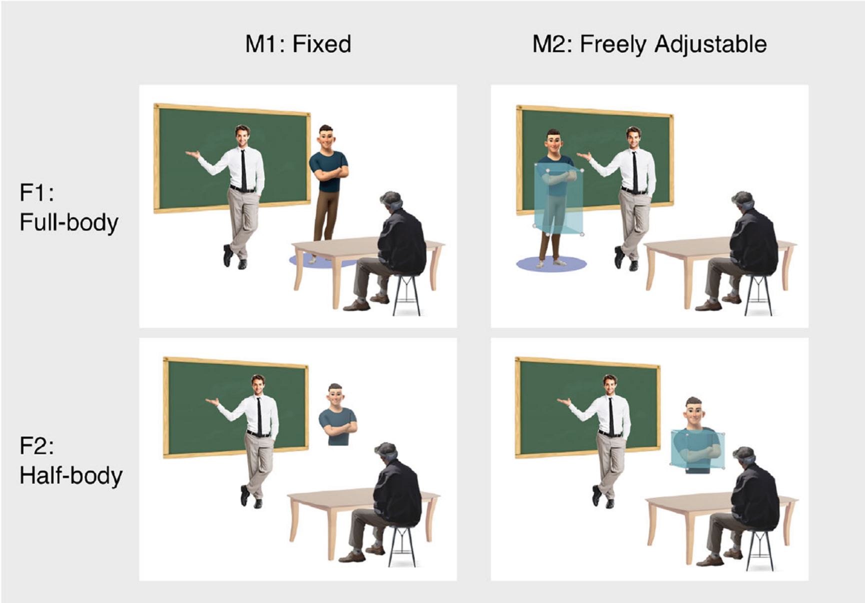


FIGURE 5 The four experimental conditions of our study: M1×F1—full body avatar with fixed position; M2×F1—full body avatar with user-controlled position, orientation, and scaling; M1×F2—half body avatar with fixed position; and M2×F2—half body avatar with user-controlled position, orientation, and scaling

into insightful research results.[21] We identified and categorized the most spoken narrations, performed interactions, and expressed feelings from the study participants.

Last, we measured the placement of the ASL avatar among each participant through the recorded video from the head-mounted display (HMD). We observed their interactions and compared them with their answers to the survey.

4.4 | Procedure

A recreated classroom environment included a speaking instructor and a participant familiar with ASL. The instructor stood in front of the class, teaching a math lesson. At the same time, the participant sat in the class, wearing the MR headset, and learning by looking at the holographic signing interpreter. After the instructor started teaching, the research team manually played the signing avatar animation.

Every user testing session took around an hour. Each involved a student helper (referred to as the "instructor") and a study participant (see Figure 6). When the study began, participants were introduced to the MR headset and given time to get familiar with the device. We first asked the participants to run the Eye Calibration through system settings and taught them to toggle the home menu by pointing at their wrists. We then asked them to open the Holographic Remoting application on the APP menu to run the Unity application. The research conductor could see their interactions through the Holographic Remoting plugin in Unity to monitor and give instructions. We then asked them to click on the Recording button from the MR headset home menu to record and document all interactions viewed from the headset. In addition, we set up a camera on the side to record the whole process from a third-person viewpoint.

Participants were asked to go through the hand gesture tutorial and get familiar with manipulating the signing avatar. The tutorial showed the main hand gestures one needed to perform near select, air tap select, repositioning, rotating, and scaling holographic objects in the MR space. Please see Figure 7 for the hand gesture tutorial.

Afterward, the math lesson began. Since none of the recruited participants were deaf, we provided earplugs to participants so they would not be distracted by the speaking instructor. We asked the instructor to read through the lecture script we provided and mimic teaching the course content. There were four math lessons, each lasting around 2–3 min; each was used to test the four experimental conditions (M1×F1, M1×F2, M2×F1, M2×F2). In the experiment, we used a think-aloud protocol, where participants were encouraged to say what they felt or thought during each experimental condition. After each experimental condition, we asked the participants to self-report and leave comments about their experiences.

FIGURE 6 Experimental setup of our study

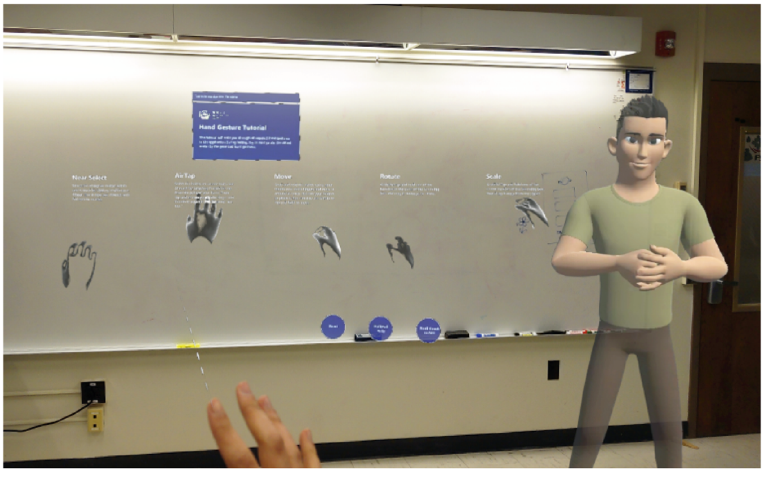


FIGURE 7 Hand gesture tutorial

After testing all four experimental conditions, we asked the participants to fill out a survey to evaluate their experiences and provide their ratings on different research measurements. Last, we conducted open-ended interviews which were recorded and transcribed for qualitative coding analysis through Otter.ai.[1]

5 | RESULTS

We separated our results into survey responses, feedback from the open-ended interview, and observations during the testing sessions.

5.1 | Survey data analysis

In our survey data analysis, we examined comfortability of the MR headset, evaluation of hand gesture tutorial scene, evaluation of avatar manipulation, comparison between avatar manipulation and avatar framing experimental conditions, and overall experience, and results from NTLX.

[1]https://otter.ai/

5.1.1 | Comfortability and acceptability of MR headset

None of the participants paused nor withdrew from the study due to motion sickness from the headset. But P2 and P3 mentioned they felt a bit dizzy after wearing the headset for more than half an hour. P2: "I'd say I did have a little bit of headache and dizziness after using it." and P3: "It gets a bit too tight and makes me dizzy." P5 mentioned the weight of the headset distributed evenly: "I used Oculus before, and I feel like the weight of this headset was distributed very well comparing to that." P1 mentioned that it felt heavier after wearing it for a while: "I think over time it got a bit heavy, and it's a lot of weight on your head. But it's okay for a short period of time." P6 think the HMD is heavy for kids "This thing is too heavy for kids. High school or college students might be fine."

The device's restricted field of view (FoV) also makes it a bit difficult to look around and navigate the holograms. P5 mentioned "With the cut offs, it feels kind of weird because I need to look up or look in a different direction to find things."

A few participants mentioned the social and psychological barriers and concerns of adopting the HoloLens HMD in daily activities. For example, P4 mentioned "Deaf people already feel ostracized from the society, adding this extra thing might make them feel even more excluded. I don't even want to use it myself in class with other students." and P5 said "I feel like if someone were deaf and they use this in a classroom, they might feel a little weird having a big headset on them. And they might be discouraged because they don't want to look strange. It'll be perfect if it's made like a pair of glasses." P6 said "I feel like little kids would just rip this thing off because it might be annoying to them, and little kids do not sit still. I didn't even want to wear glasses when I was small, I thought it was lame." P8 mentioned: "My concern is that a DHH student is already singled out sometimes, so having this headset on might be discouraging for them to use. An airer device, like glasses, might be better for them."

5.1.2 | Comments on hand gesture tutorial scene

Six study participants thought that the hand gesture tutorial scene was very useful for them to get familiar with the interactions of the application and rated seven for (HG1), while two rated four. P6: "I like the tutorial. If we just started to the lesson I would have been like, what's going on here. I like looking at it and give myself a little bit of time to try out the gestures on the menu.", P8: "Ya, the tutorial helps me figure out how to maneuver.", and P4: "It's okay. I can see it's essential, but I still find it hard to move it around."

Two study participants went through the tutorial scene in less than 5 min, five spent around 10 min, and one spent more than 10 min. All study participants were able to Reset and Toggle the Hand Coach Menu during the tutorial. However, none of the study participants clicked on any of the buttons or tried to reset the avatar's position/orientation/scale, or open the hand coach menu during the math lectures.

5.1.3 | Manipulation of the ASL avatar

Most study participants spent less time maneuvering near select than air tap, but all performed air tap during the lectures. The reason was that none of the study participants positioned the avatar close enough to do near select, all of them preferred to position the ASL avatar beyond a grabbable distance.

Rotation appeared to be the most difficult one to operate for the study participants among all five manipulation interactions (see Figure 8 for the self-reported rating). Some participants argued that the design of selecting the rotation button before rotating was unintuitive and made it challenging to perform the task. P1: "I was struggling to rotate. It's easier to rotate by my wrist while grabbing it, but I couldn't air tap and select to rotate.", P2: "Ya, rotating is kind of difficult to use.", and P7: "I feel like resizing is definitely more tangible than rotating."

Study participants also shared their thoughts on alternatives for moving the virtual avatar that might be effortless. P1: "I think the display of box should encapsulate the whole person. And it's frustrating when you couldn't find the box right away.", P5: "I prefer if it moves where I look because that's the way I'm expecting. If I turn to something else, I know it's gonna follow. I feel like that's more convenient.", and P6: "It'll be cool if the interpreter moves when I move, no wait, that sounds terrible, because I'll be looking at him even when I don't want to."

Many mentioned that we should adjust the design of the rotation function. P6 suggested simulating the interactions on smartphones. P6: "I'm thinking about swiping on iPhone. Usually you can just swipe, and it will turn or rotate on phones." P1 mentioned having the HoloLens detect wrist orientation and rotate accordingly. P1: "If the HoloLens could

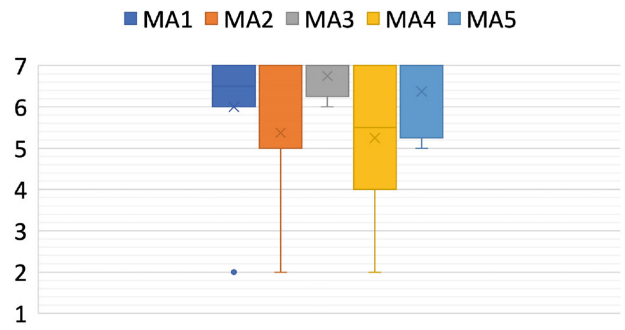


FIGURE 8 Manipulation interactions ratings. MA1: near select, MA2: air tap, MA3: move, MA4: rotate, and MA5: resize

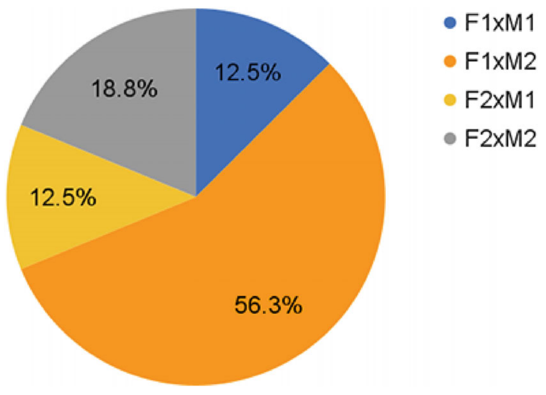


FIGURE 9 Pie chart showing the participants' selection of the preferred condition

see my wrist movement and just turn the avatar when I rotate my wrist like this, that might be a way." P2 said enlarging the rotation icon, so it is easier to select. P2: "If you can just pinch on anywhere on the side of the box, not just one spot in the middle, it'll be easier." P3 suggested to disable the natural orientation while repositioning. P3: "I think rotation should be disabled when moving, cuz sometimes it turns when I move, and I need to turn it again."

Scaling interaction was very intuitive to all study participants. P1 suggested the half-pinching hand gestures, resembling resizing an image on smartphones, as an alternative to explore. P1: "If you can just move your hand to the center of the avatar to select and scale when you do pinching, like the one you would do to resize an image on phone, that would be more intuitive to me."

5.1.4 | Comparison of the four experimental conditions

We found that full-body display with the ability to freely manipulate the avatar was the preferred experimental condition (see Figure 9 for the selection of the preferred condition). One participant preferred both full-body and half-body display with the adjustable feature, contributing .50 point to each condition.

A one-way repeated measures analysis of variance (ANOVA) was conducted to compare the four experimental conditions on participants' satisfactory ratings. There was not a significant result across the examined experimental conditions ($\Lambda = .50$, $F[3, 5] = 1.667$, $p = .288$, $\eta_p^2 = .357$). We noticed that most participants rated (see Figure 10) their experience partially on the accuracy of the ASL animation instead of entirely on the framing and manipulation settings.

All participants mentioned that the feature to be able to move the avatar is essential when needed. However, none of the participants was constantly manipulating the avatar during the lectures, even when it was overlapping with the instructor. Two participants said they could still read the signing clearly when the avatar overlapped with the instructor. Manipulation would occur only when the overlapping lasted for too long. Moreover, manipulating the avatar would be distracting while reading its signs. P6: "When the avatar overlaps with the instructor, I can still tell what he's signing, even

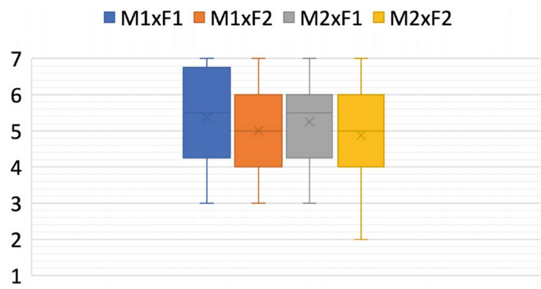


FIGURE 10 Satisfaction ratings of the four experimental conditions

though it's a little distracting, I still wouldn't want to move it cuz moving it when he's signing was even more distracting.", P4: "I understood him well enough I didn't need to move it. And moving him is very time consuming for me.". In addition, the participants also reported P2: "I like it fixed during the lecture but would prefer being able to move it when someone's moving around and gets in the way of the avatar.", P3: "If the instructor is using the board, and the avatar is overlapping with what's on the board, then I would want to be able to move it.", P6: "I am familiar with looking at an interpreter, and a real interpreter doesn't move around. So, I would say I prefer it being fixed at a place.", and P5: "I didn't move it during the lectures because I didn't want to distract myself from the signing."

Four participants preferred full-body display, three preferred half-body display, and one had no comment on different framing methods. P2 who preferred full-body display, mentioned that it looked more human-like and more natural, P2: "Full-body looks more human-like to me, half body is a bit weird." However, three out of the four participants who preferred full-body display also argued that there were no significant differences between these two framing methods for math lessons. P1 and P8 stated that for scenarios apart from classroom settings, utilizing the full-body display might be necessary. P1: "Full-body display gives a better image of the whole-body movement. Half-body would work for most academic environments, but if it were an alternate subject, like a music class when there's choreography involved, you'll need to see everything." P8: "Half-body is fine for math lesson but if you're talking about maybe shopping pants, you'll need the lower body." Two participants (P5 and P6) that preferred half-body display mentioned that it would be nice to save some space in case the avatar overlaps with more real objects. P5: "It feels convenient to have it displayed in half-body, because you don't need to worry about it being lined up with objects behind." and P6: "I feel like you can see more things if it's half-body." P7 mentioned that half-body display makes them focused on the signing. P7: "I think full body might be more distracting. I prefer half-body. I feel like I can focus more on his sign and gestures." P4 also said "I don't have a preference. But I also don't need to know what his legs are doing." P5 stated that due to the FoV of the device, there was not much difference. P5: "There's not that much of a difference since it cut off at his torso."

5.1.5 | Overall experience and NLTX results

Most test subjects stated that the application is fun to use and very refreshing. The intriguing element of such MR application for classroom settings might be beneficial for DHH students or hearing students learning ASL. Most participants could perform actions they desired according to the result of (OE1). And for (OE2), none of the test subjects rated their acceptance rate of the application as 7 due to the lack of accuracy on the ASL animation and the challenge to rotate the avatar intuitively. Please see Figure 11 for the result of (OE1) and (OE2).

Results from the NLTX form (see Figure 12) show that the mental demand (M), temporal demand (T), and frustration (F) operating the application were relatively high, while the physical demand (P) was notably low. The results implied that first-time users' interaction with MR objects is often frustrating and time-consuming. The classroom setup did not require participants to move apart from performing hand gestures. Thus, physical demand was lower than mental demand.

5.2 | Open ended interview feedback

The first question of the interview asked whether the test subject considered this application beneficial to DHH students in class and asked for an elaboration. Six participants gave positive feedback on the utilization of such technology. P1:

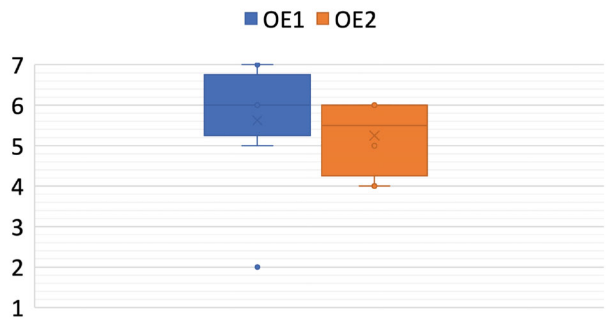


FIGURE 11 Overall experience ratings

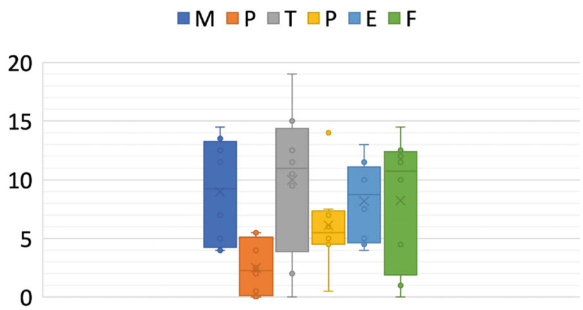


FIGURE 12 NLTX results. M: mental demand, P: physical demand, T: temporal demand, P: performance, E: effort, and F: frustration

"Yes, if there're several DHH students and only one interpreter, this could be very helpful.", P2: "Yes, I feel like having an interpreter might be intimidating for some students and isolated them apart from the hearing students. I think this technology might be a cool new way of doing things. Given that you don't need to rely on an interpreter all the time.", P3: "Yes, especially when you don't need to pay and find different interpreters for different scenarios, and the technology can just translate it for you.", P6: "Yes, for main-stream kids. But I could see interpreters being unhappy losing their jobs. Ya, I could see it being useful for high school or college students.", P7: "Yes, it could be an alternative if there's no interpreters around and the professor doesn't know sign language.", and P8: "Yes, interpreting service would be more available and affordable.". P5 mentioned the lack of two-way communication ability "It depends, because if the DHH student wants to ask a question, it'll be difficult for the instructor to know." and P4 stated that this application could not cater to the needs of DHH students. P4: "No, I think there are way too many environmental and situational dependencies that current technology could not cater. This device does not provide the flexibility a human interpreter can provide."

The second question asked about features that might enhance this technology and why. Some mentioned the importance of two-way communication, especially in classroom settings. P8: "It's great to receive the information, but you might need the feature to release information and understand by hearing people that don't know ASL." The holographic ASL agent should communicate, translate, repeat, or elaborate on the course content; otherwise, it could not replace a real interpreter. P1, P2, P5, and P7 mentioned that it would be more acceptable if the HMD were more portable. P4 said the additional feature to record or replay after class would be a nice feature "Maybe having a feature to rewind the animation would be nice. So the student can look at it again after class." P6 mentioned the possibility to add real-time subtitles that capture the audio from the speaking instructor and display it right away on a holographic board, which resembles the study of Holographic Speech Bubbles.[12-14] P6: "I think a caption box could be a nice feature. I'm not saying we don't need the digital interpreter, because having the interpreter is obviously more engaging compared to reading text from the box. But say if I missed something he's signing, I could just quickly read the caption to catch up."

The third question asked what other scenarios might benefit from adopting this application. Four study participants mentioned the possibility of adopting this application in daily conversation with hearing groups, going to the courthouse, or looking at news at a sports bar. P5 and P6 study participants mentioned the possibility of using it for entertainment purposes. P5: "Maybe a movie, and like not to worry about not having subtitles." and P6: "I could think of a million things. Deaf people's activities are limited. They could go to movie theaters, plays, or comedian shows with this." The main reason is that finding interpreters is usually time-consuming and costly. It is not practical for DHH people to hire interpreters for every activity. Such technology might provide possibilities for them to participate without scheduling an interpreter. P5: "I feel like this could help deaf community who want to do more hearing things. They don't have to hire an interpreter all the time." P1, P3, and P4 mentioned that such an application can be utilized for ASL learning purposes, which can benefit young deaf children or ASL beginners. For example, a student learning ASL through an MR headset, practice to sign while looking at the signing avatar, and the system detects the student's hand gestures to evaluate their accuracy. It could be beneficial to hearing community learning ASL as well. P1: "It would be cool if we use this in our ASL class. I can see us using this to practice signing ourselves. You can make it as an assignment, just follow what the person is signing.", P3: "There could be digital slides or something to learn the signs.", and P4: "It's functional as a learning tool, it's useful for ASL learners who aren't Deaf. And you can practice signing without a in-person teacher."

5.3 | Observations from video recordings

Through the recorded videos we observed that three participants scaled the avatar smaller than normal human size, while other remained the size of a human. Nonetheless, all of them stated that they preferred the avatar to be displayed matching the size of a normal human being. P1: "If it's doing the role of an interpreter, I would like it to feel like an actual person." In addition, we noticed that none of the study participants preferred to place the avatar within arm's distance. Thus, no near select interaction was detected during the math lessons. Most study participants preferred to position the avatar in front of the classroom while a gap remains between the instructor and the ASL avatar. We found that seven participants did not change the position of the avatar between each lecture, although they were told they could. Only one of the participants changed the avatar from his left to his right after three lessons. In Figure 13, we provide our participants' ASL avatar placement choices. The blue FoV indicates the restricted area one can see without head movements. Participants that placed the avatar closer to the instructor might be looking in another direction.

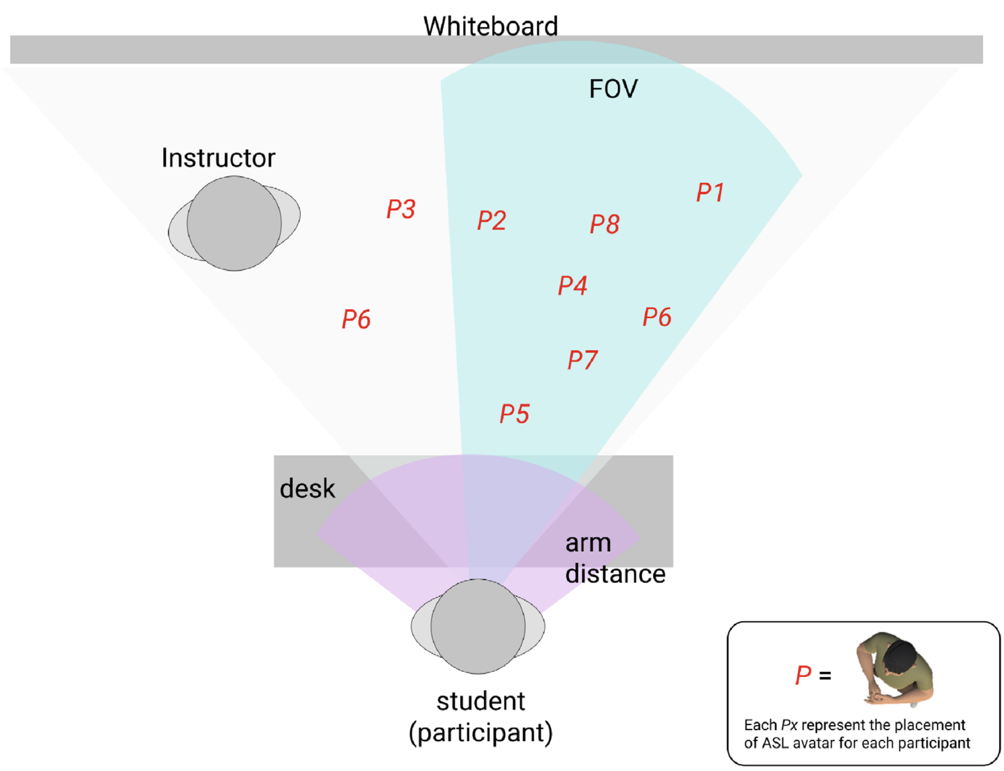


FIGURE 13 Placement summary of the ASL avatar

6 | DISCUSSION

6.1 | Holographic ASL interpreter manipulation and framing design

According to the results from the testing sessions, most study participants preferred the ability to reposition and rotate the ASL avatar but would not be performing these interactions when the avatar was signing. Study participants would only move and rotate the avatar during natural pauses and breaks within the lecture while the holographic interpreter rests and stops signing. We suggest removing the resizing feature in future development since most users prefer the size of the ASL avatar remain humanoid. Although they were aware of the cartoonish appearance of the holographic avatar, viewing it in MR space provided a robust human factor to the experience. It made the participants prefer scaling the avatar to match the size of the in-person instructor. Moreover, we discerned that the manipulation feature empowered users to personalize the experience better.

Four participants preferred the full-body display, three chose the half-body display, and one did not have particular preferences toward the two display methods. After going through their rationale, it was clear that the answer to the question on user preferences toward the two proposed framing methods would depend on different scenarios. In classroom setups, the half-body display could most likely be an option for users who prefer not to overlay the avatar with too many real-world objects. But for scenarios where the full-body movement of the signing interpreter is essential to the viewers, a full-body display must be the default display. Thus, the proposed framing design would be making the full-body display as default and remaining the half-body display an option for users.

6.2 | User tendencies

As we covered briefly in the previous sections, most participants familiar with using ASL in their daily lives or reading signs tend to regulate the virtual ASL avatar so that it remains human-like. In addition, manipulation interactions towards a signing avatar should be as intuitive as possible. We observed that the default positioning, orienting, and scaling widgets provided by Microsoft MRTK were not intuitive enough for the users to manipulate while reading the signs simultaneously. The current interactions for all three manipulation functions are composed of two actions each, selecting and dragging. A new design that simplifies the required actions from two steps to one for all three manipulation functions would be an ideal enhancement of the interaction design for the application.

6.3 | User acceptability

The participants expressed concern about the inconvenience and alienated feelings caused by wearing a bulky HMD in class. Even though it was intriguing to operate and fun to try on the device for a user study, the practicability for DHH students to wear the HoloLens HMD to class is relatively low due to physical and psychological concerns. Most people cannot wear the HoloLens device for an extended period because they might get motion sick from the HMD. HoloLens appears to be better than most VR HMDs in this aspect, given that the weight of the headset is distributed more evenly, and users can still navigate the real environment through the glasses. But with the heaviness of the device, users might feel intense discomfort in their neck after using it for a while. Let alone sitting in lectures that might last 2–3 h.

The requirement for the system to perform two-way communication translations is also a critical factor in boosting the acceptance rate of this application. The significant element in-person interpreters have, which digital devices could not thoroughly replace, is that they provide immediate alternations and solutions that cater to a DHH person's need. For example, when a DHH person requires an elaboration of a particular interpretation or when a DHH student asks questions in class, the in-person interpreter can cater to these requests while the digital ASL avatar on the HMD device cannot. However, this is just feedback regarding the current limitation of technology, which we think researchers could resolve in the future.

6.4 | Limitations

The biggest limitation of our study is that there were no DHH individuals who signed up for the testing sessions. Although two were native ASL users, and two had used ASL for around 10 years, most of the participants were used to naturally engaging with speaking instructor in class through listening to them teaching. Moreover, we consider the sample size of our study as an additional limitation. Nevertheless, this study provides the foundation for future user-centered research on holographic sign language avatars in MR space.

7 | CONCLUSIONS AND FUTURE WORK

The results of this study show that the scalability of computer-generated sign interpreters allowed users to configure settings freely according to their preferences. At the same time, sign language users prefer the MR interpreter to resemble a natural person as much as possible.

In our future work, we plan to conduct a study with DHH students in mainstream schools and enhance the application so it can cater to different classroom scenarios. In future implementations of the application, we plan to integrate Azure Spatial Anchors in order for the system to scan the environment and make sure it places the avatar in the correct starting location. We also plan to implement speech bubbles, which we consider a promising add-on feature for classroom settings or other scenarios. Moreover, according to the feedback from the participants, this application can also be used as ASL learning tool for both hearing and deaf communities. Thus, we would like to also implement features such as hand gesture detection and evaluation to rate the performance of the students. An MR ASL learning application could help bridge the gap between hearing and DHH communities by enhancing the learning efficacy of ASL learners.

ACKNOWLEDGMENTS

This research project was funded by Purdue University, Instructional Innovation Grant.

ORCID

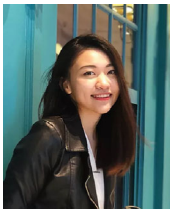
Fu-Chia Yang https://orcid.org/0000-0003-2041-4836
Christos Mousas https://orcid.org/0000-0003-0955-7959

REFERENCES

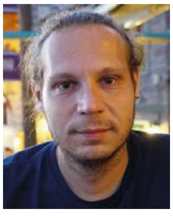
1. Constantinou V, Ioannou A, Klironomos I, Antona M, Stephanidis C. Technology support for the inclusion of deaf students in mainstream schools: a summary of research from 2007 to 2017. Universal Access Inf Soc. 2020;19(1):195–200.
2. Zirzow NK. Technology use by teachers of deaf and hard-of-hearing students. Morgantown, WV: West Virginia University; 2019.
3. Saladin SP. Psychosocial variables in the adoption of assistive technology among deaf and hard of hearing adults. Austin, TX: The University of Texas at Austin; 2004.
4. Huenerfauth M, Hanson V. Sign language in the interface: access for deaf signers. Universal access. Volume 38. Handbook NJ: Erlbaum; 2009. p. 14.
5. Kipp M, Heloir A, Nguyen Q. Sign language avatars: animation and comprehensibility. Proceedings of the International Workshop on Intelligent Virtual Agents. New York, NY: Springer; 2011. p. 113–26.
6. Kaneko H, Hamaguchi N, Doke M, Inoue S. Sign language animation using TVML. Proceedings of the ACM SIGGRAPH Conference on Virtual-Reality Continuum and its Applications in Industry. New York, NY: Springer; 2010. p. 289–92.
7. Lee J, Kunii TL. Computer animated visual translation from natural language to sign language. J Visual Comput Animat. 1993;4(2): 63–78.
8. Kaur K, Kumar P. HamNoSys to SiGML conversion system for sign language automation. Proc Comput Sci. 2016;89:794–803.
9. Kaur R, Kumar P. HamNoSys generation system for sign language. Proceedings of the International Conference on Advances in Computing, Communications and Informatics, Delhi India. IEEE; 2014. p. 2727–34.
10. Kang Z. Spoken language to sign language translation system based on HamNoSys. Proceedings of the International Symposium on Signal Processing Systems, Beijing, China; 2019. p. 159–64.
11. Vesel J, Robillard T. Teaching mathematics vocabulary with an interactive signing math dictionary. J Res Technol Educ. 2013;45(4):361–89.
12. Peng YH, Hsi MW, Taele P, Lin TY, Lai PE, Hsu L, et al. Speechbubbles: enhancing captioning experiences for deaf and hard-of-hearing people in group conversations. Proceedings of the ACM CHI Conference on Human Factors in Computing Systems. New York, NY: Springer; 2018. p. 1–10.
13. Jain D, Franz R, Findlater L, Cannon J, Kushalnagar R, Froehlich J. Towards accessible conversations in a mobile context for people who are deaf and hard of hearing. Proceedings of the ACM SIGACCESS Conference on Computers and Accessibility. New York, NY: Springer; 2018. p. 81–92.
14. Guo R, Yang Y, Kuang J, Bin X, Jain D, Goodman S, et al. HoloSound: combining speech and sound identification for deaf or hard of hearing users on a head-mounted display. Proceedings of the ACM SIGACCESS Conference on Computers and Accessibility. New York, NY: Springer; 2020. p. 1–4.

15. Vinayagamoorthy V, Glancy M, Debenham P, Bruce A, Ziegler C, Schäffer R. Personalising the TV experience with augmented reality technology: Synchronised sign language interpretation. Proceedings of the ACM International Conference on Interactive Experiences for TV and Online Video. New York, NY: Springer; 2018. p. 179–84.
16. Vinayagamoorthy V, Glancy M, Ziegler C, Schäffer R. Personalising the TV experience using augmented reality: an exploratory study on delivering synchronised sign language interpretation. Proceedings of the ACM CHI Conference on Human Factors in Computing Systems. New York, NY: Springer; 2019. p. 1–12.
17. Adamo-Villani N, Anasingaraju S. Holographic signing avatars for deaf education. E-learning, E-education, and online training. New York, NY: Springer; 2017. p. 54–61.
18. Miller A, Malasig J, Castro B, Hanson VL, Nicolau H, Brandão A. The use of smart glasses for lecture comprehension by deaf and hard of hearing students. Proceedings of the ACM CHI Conference Extended Abstracts on Human Factors in Computing Systems. New York, NY: Springer; 2017. p. 1909–15.
19. Parton SB. Glass vision 3D: digital discovery for the deaf. TechTrends. 2017;61(2):141–6.
20. Hart SG. NASA-task load index (NASA-TLX); 20 years later. Proceedings of the Human Factors and Ergonomics Society Annual Meeting; Los Angeles, CA: Sage Publications; Vol. 50, 2006. p. 904–8.
21. Smith CP. Content analysis and narrative analysis. Handbook of research methods in social and personality psychology. New York, NY: Cambridge University Press; Vol. 2000, 2000. p. 313–35.

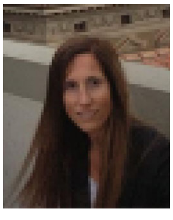
AUTHOR BIOGRAPHIES



Fu-Chia Yang is an MS student at Purdue University majoring in Computer Graphics Technology. She graduated from The Chinese University of Hong Kong with BS degree in Computer Science. Her research interests include virtual reality, augmented reality, character animation, and human-computer interaction.



Christos Mousas is an assistant professor at Purdue University and director of the Virtual Reality Lab. His research revolves around virtual reality, virtual humans, computer animation, applied perception in computer graphics, and immersive interaction. From 2015 to 2016, he was a postdoctoral researcher at the Department of Computer Science at Dartmouth College. He holds a PhD in Informatics and an MSc in Multimedia Applications and Virtual Environments both from the School of Engineering and Informatics of the University of Sussex, and an integrated master's degree in Audiovisual Science and Art from Ionian University. He is a member of ACM and IEEE, and has been a member of the organizing and program committees of many conferences in the virtual reality, computer graphics/animation, and human-computer interaction fields.



Nicoletta Adamo is a professor of Computer Graphics Technology and Purdue University Faculty Scholar. She is an award-winning animator and graphic designer and creator of several 2D and 3D animations that aired on national television. Her area of expertise is in character animation and character design and her research interests focus on the application of 3D animation technology to education, human computer communication (HCC), and visualization. She is co-founder and director of the IDEA Laboratory.

How to cite this article: Yang F-C, Mousas C, Adamo N. Holographic sign language avatar interpreter: A user interaction study in a mixed reality classroom. Comput Anim Virtual Worlds. 2022;e2082. https://doi.org/10.1002/cav.2082