Designing for Social Interactions in a Virtual Art Gallery

Nicholas F. Polys
Virginia Tech
USA
npolys@vt.edu

Samridhi Roshan
Virginia Tech
USA
samridhi18@vt.edu

Emily Newton
Virginia Tech
USA
emilyrn18@vt.edu

Muskaan Narula
Virginia Tech
USA
muskaannarula22@vt.edu

Bao-Tran Thai
Virginia Tech
USA
tranthai@vt.edu

Figure 1: This story board (comic) shows two friends using the virtual art exploration feature to find exhibits around the world based on location. This is an example of how users can interact with the application to discover different interests and exhibits together. Drawing by Muskaan Narula.

ABSTRACT

The dawn of a new digital world has emerged with new ways to communicate and collaborate with other people across the globe. Metaverses and Mirror Worlds have broadened our perspectives on the ways we can utilize 3D virtual environments. A Mirror World is a 3D virtual space that depicts a real-life place or environment that people may want to see physically or would like to manipulate to create something new. A perfect example of this would be an art gallery which provides people an outlet to express themselves through various art forms and be able to socialize and have that human interaction that is needed during times when physical presence may be difficult.

This project strives to improve user social interactions and make spatial control easier and more fluid in a virtual art gallery, while also incorporating the existing metaphor of permission and user privileges used in synchronous collaborative environments. We worked to create ways for people to be invited into group chats based on proximity, allowing users to give their consent as to who they want to talk to and who they will allow to share control within the space. We also implemented a way to view the space as a 3D map that highlights pieces of artwork around the space for people to teleport to and view at ease. To demonstrate this shared viewing and navigation experience we also focused on incorporating audio and spatial interaction features within the art gallery prototype of X3D and glTF models, images and audio, and HTML user interface.

Web3D '22, November 2–4, 2022, Evry-Courcouronnes, France

ACM ISBN 978-1-4503-9914-2/22/11.
https://doi.org/10.1145/3564533.3564562

CCS CONCEPTS

• **Human-centered computing** → **Visualization**; **Collaborative and social computing design and evaluation methods**.

KEYWORDS

CSCW, X3D, Usability, Metaverse

ACM Reference Format:
Nicholas F. Polys, Samridhi Roshan, Emily Newton, Muskaan Narula, and Bao-Tran Thai. 2022. Designing for Social Interactions in a Virtual Art Gallery. In *The 27th International Conference on 3D Web Technology (Web3D '22), November 2–4, 2022, Evry-Courcouronnes, France.* ACM, New York, NY, USA, 9 pages. https://doi.org/10.1145/3564533.3564562

1 INTRODUCTION

We set out to create a User Interface that can be used in art visits, art display, and attending events or shows in online virtual galleries, despite physical barriers to collaborative art viewing. Our target users are any art lover, especially the ones that cannot physically attend a show or an event at a gallery, perhaps because it is abroad or too far. Our goal is to solve the ongoing problem of user control and interactions with others in the virtual space. To do this, we will need to gain a better understanding on how people interact with one another in the physical space, so we can replicate these human ways of conveying information in a digital format. In the real world there is a lot more information that is conveyed through body language, hand gestures, and tone of voice than what is actually being said. All of this information is lost in a virtual world, and it is important that we find a way to bring this context to conversations and interactions to prevent misinterpretation and confusion between people in the mirror world.

Through this Senior Capstone project, we ran a full Human-Computer Interaction design cycle from requirements engineering to design and evaluation. We used a combination of contextual inquiry (real-world observation), a user survey, and scenario-based designs to consider a variety of backgrounds and motivations of users. To understand the nature of the design challenge, we observed real user behavior in our campus art gallery; we took notes and then using Content Analysis techniques, we extracted key themes and motifs of user activities. We then used an online survey to further develop an understanding of user perceptions toward multi-user 3D collaboration. We used this data to fine-tune and prioritize the requirements we would design for. The Design phase included several ideation sessions where we generated a seat of features to meet our requirements. We then reduced this set to be feasible in our timeframe (6 week prototyping phase). We reviewed the system with both expert and end users, deriving new insights as to how such (virtual) social interactions can be realized.

2 BACKGROUND

We have found several prior works about virtual galleries in the Metaverse. One of the examples of an existing virtual gallery is the three dimensional graphic digital construction in the RMIT University's virtual campus in Second Life where users take on a virtual avatar form to view the gallery[Anderson et al. 2013]. The authors examine different ways of interaction like augmented reality using power points, virtual environments using Second Life on a screen, and immersive environments using Second Life on a head mounted display. They found that users who used the artwork directly using powerpoint had the easiest experience. However, walking around with their avatars adds to that experience in a whole new way. Therefore, we need to find a way to provide users with an "easy metaphor" experience in the Metaverse type environment.

In a similar way, the article "Social Immersive Media," brings up another solution on how to make user communication easier in a museum setting as well [Snibbe and Raffle 2009]. They talk about the experiences and studies being conducted about using interactive camera/projector systems with social immersive media. This is a form of augmented reality that allows people to clearly express, "strong emotional responses and social engagement through visceral interaction," (Ibid, p. 1447). They also discuss how they hope to use this technology in science, history, and art museums so users can be more engaged with the space and items since their whole body would be placed in this synthetic mirror world. Some examples of these emotional responses would be a visitor jumping with excitement or the physical action of laughing out loud in a place where, in the real world, it would be uncommon. It also implies that it may be necessary for people to over exaggerate their emotions in a digital world to get their point across to others. In the paper, "Virtual studio practices", "Working virtually offers a range of interesting benefits for creative practice" [Budge 2013]. However, according to the paper, the current virtual studio practices work lacks the face-to-face interaction and the materials sharing approach. This further supports our notion that we need to think of ways to exaggerate facial expressions and body language; this is a worthy animation challenge in and of itself.

Kim's "Social TV Viewing" refers to the phenomenon of viewing TV all the while communicating with others privately, or publicly about it [Kim et al. 2021]. This includes using applications such as Netflix Teleparty, or watching live TV all the while tweeting/ texting/ video chatting/ voice calling your friends about it. Communicating with one another while watching the same content in real time results in the feeling of "social presence." The study found that this kind of social presence is an important way to connect with those one cannot be physically present with. It can increase engagement of those who are on the platform, and aims to facilitate channels with friends.

The Metaverse can be persistent and also a constantly updated collection of mixed reality spaces mapped to different spatial locations for mobile mixed reality applications. Navigation among multiple virtual spaces and providing enriched spatial context is a challenge. In the past, some major issues of such applications include not allowing users to traverse among multiple active spaces and the lack of remote collaboration allowing for interaction across enriched spatial contexts. However, according Aizu, Decentraland, a 3D virtual world browser-based platform, resolved these challenges [Ryskeldiev et al. 2018].

Mixed reality applications use virtual spaces where both local and remote users can collaborate and such applications support both multiple collaborative spaces in a session allowing users to see all available spaces sorted by locations and select them in a list or as icons of spaces representing a user's current location based on geographical coordinates of the space they are in. Users in a single space can collaborate together in real time through audio and video streaming as well as 3D annotations in virtual space. The whole Metaverse is stored in a JSON array and can be shared in plain text form. As a future possibility, collaborative spaces could be accessed using a QR code in public spaces.

User participation in computer game modifications and Massively Multiplayer Online (MMOs) have grown over the years with

a top-down development style. 'Gamifying' a virtual space can lead to longer and more meaningful user engagement. Volk [Volk 2008] describes the approach of game development that intertwines multiple platforms and their concepts of development. The trends that merge a collateral tension between conventional development and "prod-usage" are not just specific to characteristic changes towards the Metaverse.

Ahn et al [Ahn et al. 2001] discuss how they created a VR theater in Korea that gives people a virtual experience of an IMAX theater. They explain how this multi-user space has some issues on how the audience can interact with the theater and others around them. This theater is meant to, "provide a visual, aural and olfactory immersive system for a large audience," and it is supposed to allow over 600 people to interact with the video being shown. Some of the issues they faced is how the multi-user interactions have to be rendered in real-time, but adding these interactions into a movie that has a fixed story in a real IMAX theater is something that is difficult to accomplish. They also discuss how this VR theater is being edited by hundreds of people all on the same screen, which may be overwhelming and they even say is something that needs to be reconsidered when attempting to create a large scale multi-user platform in VR.

The authors also introduced a concept called, inhabited TV, where multiple users could interact and participate in the TV show. This is where they introduce a possible solution to multi-user interactions on the same screen called layered participation. This invites the idea of having a hierarchy where some users have more or less privileges for how they interact with the TV show. This article gives us some insight on some of the issues we may face when handling a multi-user environment in the mirror world. It also gives us some solutions that have been researched and tested, so we can make educated decisions on what our art and performance viewing mirror world platform will want to include in the design to ensure the best multi-user experience [Ahn et al. 2001] (p.41).

It is also important to understand how social presence is created in a more physical, spatial context. Banayan (2022) considers the factors that go into creating a gallery in a museum. Curators look for different and creative ways to combine spatial elements and have them positively affect the art. Museums no longer separate art by mediums but by other ordering methods like chronology (Banayan, 2022). The article mentions that MoMa even lets people design their own viewing path often (Banayan, 2022). Museums must also weave in transparent, empty spaces to encourage reflection and engage conversation amongst people, stating the viewers must "feel like an active agent, not a passive one" (Banayan, 2022). This is critically important to our work as it helps us consider the social and spatial factors that we may have to consider for the virtual art platform we are creating.

3 REQUIREMENTS

3.1 Observations

Our team decided to conduct observations at the real, physical Moss Arts Center (MAC) to get an understanding of how people interact with a physical space that was previously used for cross-platform Mirror World installations [Polys et al. 2015]. We chose to separate these observations out so we could get different perspectives based on the day of the week. The days and times we chose were: Wednesday 3pm, Thursday 2PM, and Friday 12pm. Our observations on the real space can be summarized as follows:

- A minimalistic architecture with a neutral color scheme of whites and grays much like a canvas. These colors and the high vaulted ceilings give this building a spacious and open environment. The large and abundant windows let in a surplus of light that also weighs into this clean and calming atmosphere. There are also many windows that are on the inside of the building that allow people to look into exhibits and labs from above, which also opens up the space since it gives people the illusion of connectivity in which the hallway is part of multiple rooms and spaces of the building without the barriers of walls.
- Another attribute this space creates is atmosphere through sound. Throughout the building there are speakers that play calming music and during one of our observations they were playing jazz music that could easily be heard, but soft enough to introduce the building's character. This light audio may be a feature of importance in a virtual space as well, to set the mood and atmosphere for a user.
- The layout of the building can be confusing. There are very few signs and directional outlets for people to find their way around and it is easy to get lost
- We observed that there was one supervisor for each primary art exhibit who sat at a table near the entrance and close by to a table filled with pamphlets and information about the exhibit; when asked questions about the exhibit they were supervising, they had limited knowledge about the artists and their work.
- We noticed many people were either walking through by themselves or with one other person. There were only a handful of people who passed through to view the exhibits, and many of these people were quiet; if they were with another person they spoke in whispers while they were in the exhibits.

3.2 Survey

We seek to understand the current systems in place which people use to connect in a virtual environment. We aimed to understand two main types of virtual platforms - virtual platforms for viewing videos (for example, Netflix Teleparty, Amazon Watchparty) as well as virtual platforms for simple connectivity (e.g., Zoom, Discord). With both platforms, we aimed to understand the standards of primary control, and people's satisfaction with these rules. Understanding how users manage attention and communication while viewing will be critical to our virtual art gallery environment.

We created an informal online survey with fourteen questions. The survey received 31 responses, of which data was consolidated. Overall, more people have used meeting platforms than viewing party platforms. With viewing parties, Netflix Teleparty was the most used, followed by Disney Groupwatch and Amazon Watchparty. Among meeting platforms, Zoom was the most used followed by Discord. The three main topics of discussion included standards of primary control, communication methods, and points of frustrations.

With Viewing Party Platforms, 66.7 percent stated that the host typically has the control and the person who organizes and starts the meeting is the host. Another person stated "we had enough trust that we felt restricting it was unnecessary." These suggested there were some mixed reviews. Most, however, stated they felt it was fair that the creator of the link had all the privileges. The major issues with these applications was the lack of connectivity / lag as well as the poor User Interface. People also rated their overall satisfaction with these platforms as "somewhat satisfied." In terms of communication modalities, people said they mostly did a separate video call or used an alternative chat or video call method. Only about half of the users noted they used the builtin text chat feature.

Everyone stated that they had used meeting platforms at one point or another. The primary reasons cited were from work/school, screen sharing and hanging out with friends, respectively. Zoom was the most used platform, followed by Discord, and then Microsoft Teams and Google Meet were tied. Here as well, people mostly felt the creator of the room should be the host and have most of the privileges. Many cited examples such as "the teacher or professor was the host," to explain the need for host privileges. Another claimed that "It prevents disruptions and ensures a more smooth workflow." This seemed to be the overwhelming opinion. Comparing chat box vs voice call, people had similar reactions and reasons for using, except people preferred voice calls for gaming more.

Our survey questions focused on the two most common analogs to collaborating virtually where some digital multimedia is the prime task: watch parties and online video conferencing platforms, both of which have seen a huge rise in usage, especially after Covid. Covid has made us realize the need and importance of such virtual collaboration tools. The most used virtual collaboration tool for watch parties seems to be Netflix and for work, it seems to be Zoom and Discord.

3.3 Watch Parties

10 out of 31 people who we surveyed had never used any of the watch parties available. Users' answers suggest that people sometimes use the "screen share" feature through Zoom or Discord. However, Netflix and other companies now understand that people have been doing this. Hence, they now block their screen or black them out when someone tries to use the "screen share" services on a different company's platform.

Most of these services have a host or privilege feature. About 66.7 percent of the people said that the control during the watch parties was with the host. Hosts are the people who organize the meetings or start the watch parties and they usually have the control privilege by default. On asking how they feel about this privilege, we understood that most of the users in small groups are okay with it because there is an unspoken trust among the members, especially among friends. Also, in large groups, the organizer was the host and it seems there was a mutual understanding about this privilege being with the user.

One of the biggest features of collaboration during these watch parties is the built-in chat feature. Netflix Teleparty and Amazon Prime Watch Party are the two most common and famous platforms for large groups with capacities of 50 and 100 participants respectively [**?**]. Both of these only support text features for communication within the platform. One might think these chat features would be the most used communication tool, but surprisingly people prefer video calls over this text feature. This indicates the importance of facial expression and tone of voice while communicating. In fact, about 42 percent also use alternative chat apps or other social media platforms for communicating with their friends.

This raises the question of why have a text feature in these watch parties? They could just use their mobile phones for communication, but what makes this text feature unpopular? Our survey suggests that not a single person thinks the chat display to be great, they think the display is neutral to somewhat satisfied. In fact, 21.1 percent of users think that these watch parties have poor UIs. No wonder, since the chat box takes up quite a bit of the screen space. This can be especially annoying if the audience is engrossed in the film or art. The question then becomes what feature do we implement that can allow the users to focus on the art when they need that attention or concentration.

Some also think these services are too costly and not useful. A part of it lies in the difficulty of access to these services because only the people who have subscriptions can participate and one of the users also mentioned that the participants have to be in the same country. Also, one of the biggest concerns with collaboration during watch parties was internet lag and bad response on every participant's end. As expected, every single person we surveyed had experienced video conferencing tools in some form or the other.

We have understood that the privilege and remote controlling power depends on the situation or context of the collaboration. If it is a school meeting, then it makes sense for the professor or teacher to have the controlling power like allowing sharing screens or creating polls on zoom, creating server channels on discord, to maintain structure in the class. In an office setting, it makes sense for seniors or the managers to have control. However, if the collaboration is among a group of peers, a hosting privilege hierarchy may be awkward. This suggests that having a situation-based privilege feature is better than a universal privilege policy. In most cases where we use these platforms, both text and voice communication have almost equal preferences. However, while gaming, users prefer audio communication. This indicated that when users are actively involved in performing certain activities (particularly using the keyboard in this case), they would want audio features for communication.

3.4 Derived Requirements

After working through our background research, observations, and survey data, we identified several important factors that would improve users experience within the virtual art gallery that led us to our nine requirements. The primary factors that we wanted to incorporate into our main requirements involved how we can make user control and interactions self explanatory and easy for most users. More specifically, we want to prioritize ways for users to navigate to different exhibits or art pieces easily and create a UI that is familiar and understandable. This requirement was derived from the observations that we made at the MAC in regards to how

difficult it was to navigate the physical environment due to the lack of signage and derived from our survey responses about watch party application UIs.

We also found user interaction with other users to be of great importance from our survey responses and research. This meant we needed to require ways that users could communicate and express themselves with one another. Based on our survey responses a combination of voice chatting and text-based chatting would be something most people would be familiar with using. We would also want to create a form of private chat or proximity chat to allow small groups to be able to communicate with one another without hindering any other viewer's experience that may not want to be included.

Prioritizing user interaction with items in real time was another factor we identified in our research. This could include audio, to create context within a space or to be used as a narration to describe an exhibit or art work. Other ways users could interact with an artwork could be to leave a reaction such as a heart or a smile emoji similar to any social media platform.

These are all core factors that create the foundation of what this virtual space is meant to provide. In summary we derived our requirements listed below from the desire for self-explanatory controls and interactions, user to user communication, and the ability for users to interact with items in the space. We can summarize with a list of our derived Requirements in Priority Order. Requirement:

(1) Information Directions: Users shall click on checkpoints within library to go to exhibits
(2) Social/Object Interaction: Users shall be able to create proximity group chats
(3) Audio: Users should be able to listen to a narration of a piece of art work and control it
(4) Exploration: Users shall be able to view a global map and be able to visit exhibits around the world
(5) Social and Object Interaction: Users shall use private chatting with other attendees
(6) Social and Object Interaction: Users shall be able to leave interactions on pieces of artwork for artists and other users to see
(7) Profile/Privilege: Artist users shall be able to see number of views and length of time people stay within exhibits that they have created
(8) Social/Object Interaction: Users shall use audio chatting instead of texting
(9) Profile/Privilege: Users shall look at their liked exhibits and art works

4 DESIGNS

After our ideation sessions and drawing out our designs for the nine requirements we listed, we narrowed it down to the top three requirements to prototype in the implementation phase. The three requirements chosen were #1, #2, and #3 since they would test some fundamental social interactions in 2D plus 3D spaces and were feasible to prototype within our time constraints. Therefore, the design we have chosen is a browser accessible art gallery that will allow users to navigate and learn about the gallery through

Figure 2: User choice of un-rigged avatars

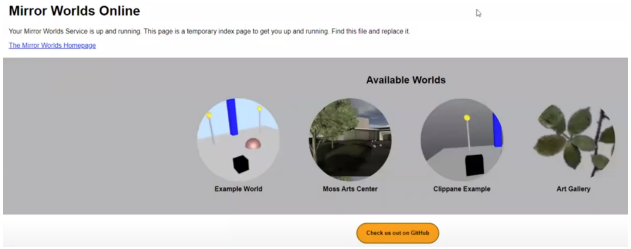

Figure 3: User choice of destination worlds (X3D and GLTF)

checkpoints on a map. This makes it easier for users to quickly teleport from one place to the other.

We will allow users to view pieces of art in a 3D space with narrations and possible animations; users will be able to control the audio for each piece. Users will be able to communicate with one another through a proximity or private chat box and/or voice chat. This chatting feature proximity specific will allow users to be able to communicate with only a small group of people to avoid a larger group chat that may become overwhelming. We developed on top of the Fusality Server software written at Virginia Tech [Polys et al. 2015]. It is a node.js service for publishing and subscribing to various events in the Mirror Worlds; updates happen through the WebSocket connection to all X3DOM clients. We created two X3D spaces to test our designs: an X3D model of our Moss Arts Center and the Sponza courtyard model. A set of 6 simple art works (images) and their audio narration were placed around the spaces. Figure 3 shows the start up screen where users pick a name, avatar, and destination world.

4.1 Prototype

In our prototype, users can navigate around in first person, or use the map view to teleport to other locations. Users can start and stop embedded media in the 3D world and can capture text notes at a location and save them for later. A text chat channel is supported as well a salient drop down list and button to invite other users to a private text chat. The user interface is shown in Figures 4 and 5. We considered information and interaction design specific to

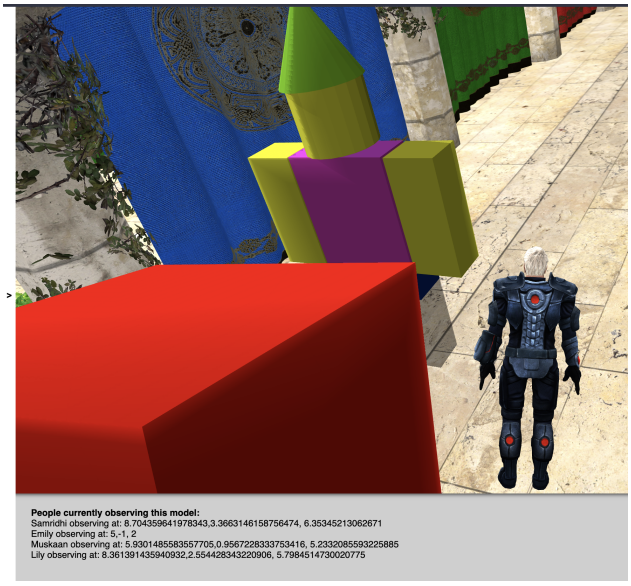

Figure 4: Screenshot with several avatars in the gallery

our five major identified requirements: Social/object interaction, profile/privilege, information/directions, audio, and exploration.

- Users shall click on checkpoints within library to go to exhibits: This feature aims at providing users an overview of the gallery as well as critical "checkpoints" they can use teleportation to reach. This enhances the navigational abilities of users as they can identify key parts of the gallery and navigate to specific artworks much more easily. When one clicks on the map icon on the bottom left corner, one can see a birds eye view of the gallery appear. Along with pins marking the "checkpoints." When the user clicks on a checkpoint, they will be teleported to that location.
- Users shall be able to create proximity group chats: At present, we have a private chat UI at the lower right corner of the screen. You can click on the "messenger looking button" at the lower right corner and it will open a pop up which contains a dropdown menu for selecting the user you want to chat privately with. The list of users are intended to be sorted based on the distance from the player user with the closest user being on the top of the list and the farthest user being at the bottom. The pop us can be closed by clicking on the button again clearly indicated by the x on the button itself.
- Users should be able to listen to a narration of a piece of artwork and control it: For each artwork, the user can have the ability to teleport to that artwork upon a single click and be able to pop up a mini screen that displays an audio menu and player as well as a miniature picture of the artwork. They also have the ability to close the artwork menu and keep the audio player open as they continue exploring the world or examining the artworks in depth.

5 EVALUATION

5.1 Quantitative Analysis

For the quantitative analysis, we performed an empirical evaluation of our product. We followed four users as they used features designed and developed in the web application. The benchmark test was a survey the users filled out in while using the application. This included questions regarding navigation within the application, the ease of use, and nature of interactions within the world. We summarize the metrics and observations below. All were asked to rate the ease of use (1-5)? Where: 1 = very difficult/ confusing; 2 = somewhat difficult/ confusing; 3 = neither confusing nor clear; 4 = somewhat easy/ clear; 5 = very easy/clear.

5.1.1 Overall Evaluation. Performing user studies made us realize the importance of metaphors in user-centric design. For most of the features we were trying to evaluate, the user was trying to connect it to some sort of already existing action or feature in technology like trying to use the arrow keys for movement, or quickly understanding the messenger button is for private chatting or having an avatar picture next to the name in the chat boxes, and so on. Overall, the most challenging task for most of our users was the navigation and movement in the Metaverse and some of the easiest things to understand were the notes and the changing pictures featured in the art/audio modals. We received quality suggestions from our users and these could be implemented in future improvements of this project. The quantitative results (median) of the four user's ratings for each of our prototype features are as follows (out of 5):

- Map/ Navigation 3.5
- Art Audio 5
- Proximity Chat 4
- Notes 4.75

5.1.2 Base Features. We aimed to quickly assess the usability of some base features. We tested users on their ability to identify and change their avatars and navigate around the world. The avatars were clear from the get-go for all users. All were able to change their avatars as well. Moving around was the difficult part. Users were confused about how to comfortably move around the world. All users tried to use the arrow keys and "WASD," realizing it did not work in this environment. Many users got lost in the world and ended up having to press "R" or refresh the page. One user was able to quickly learn how to navigate via scrolling. No one learned how to "Fly" however, which was a mode many of us liked to use when developing. One user even suggested a tutorial in the beginning would be nice.

5.1.3 Map/Navigation. The users that we asked to interact with the Metaverse collectively thought that the checkpoint on the map was a "you are here" indicator. It took more than 30 seconds for the users to figure out that the checkpoint was actually a teleporter, and it took several prompting attempts. In one testing occurrence the user experienced a bug: sometimes the navigation switched to "scroll/ orbital" motion instead of forward and backward, after a teleport occurred.

5.1.4 Art/Audio. The artwork panel that contained the audio narrations had self-explanatory buttons according to our users. For

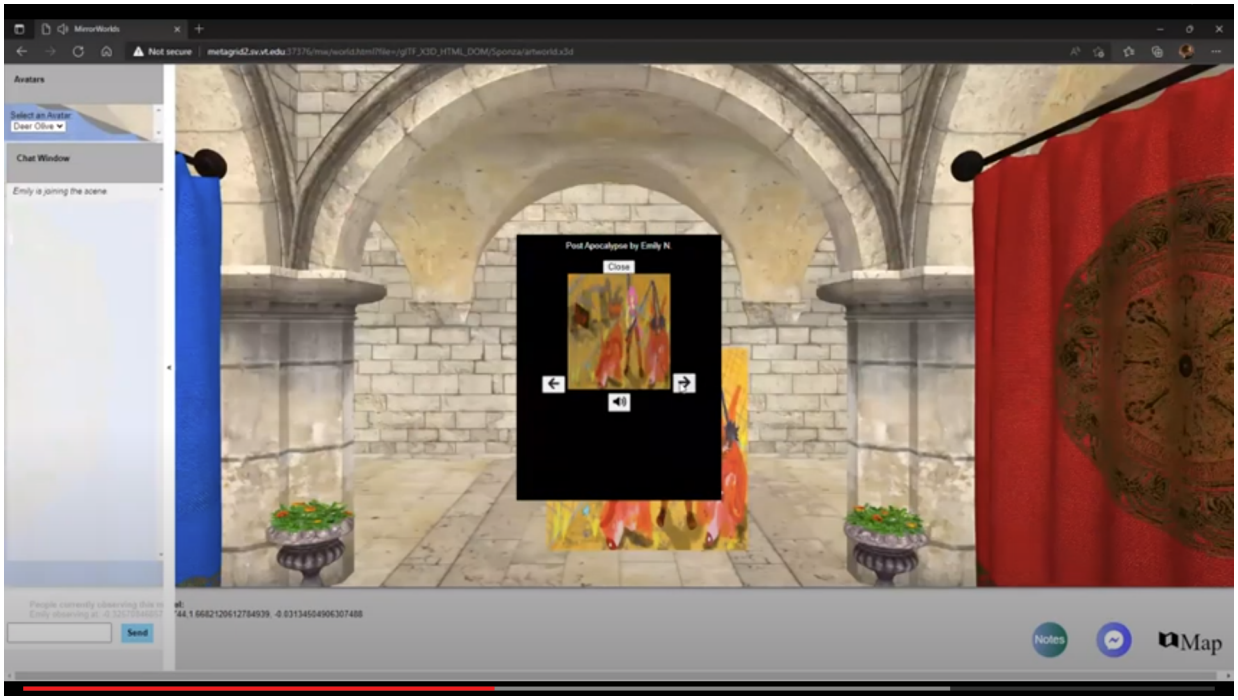

Figure 5: Private Chat dialogue with an interactive spatialized video piece

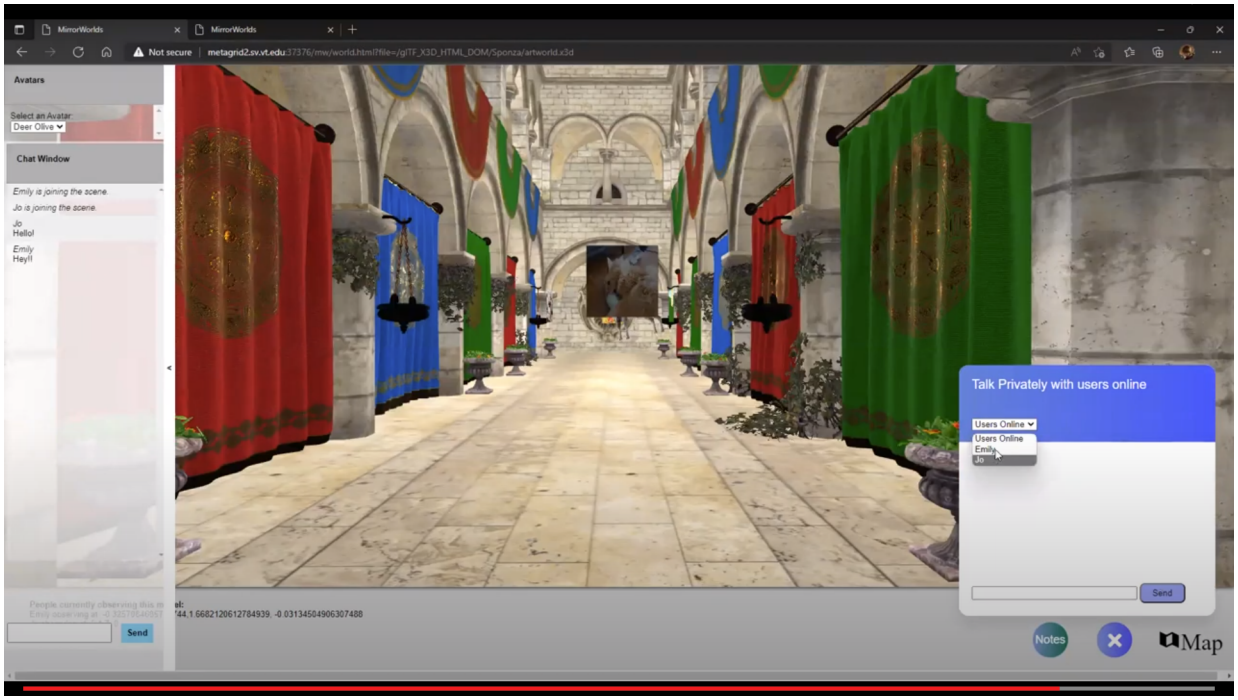

Figure 6: Private Chat dialogue in the art gallery

most of the people we had tested this the art gallery, they gravitated towards the artworks and out of curiosity, began clicking and playing around with their features. They thought the layout was understandable and the button functions were clear. A few of the critiques that we got involved the scaling of the photo in the art panel. They wished that it was larger and that they could scale it to how they want it. Another problem they found was the audio portion of the art panel does not close with the rest of the art panel and the audio continues playing even after the audio controls are closed out. The users explained that this was not what they expected to happen, and they wished that the audio controls would open and close with the art panel. However, we did get some positive feedback about how the audio affected their artwork viewing experience. One user stated that the audio feature felt like an "interactive audio tour like in a real museum," which is the type of experience we were going for when creating this feature.

5.1.5 Proximity Chat. After we asked our users to locate the public chat, one of our users thought that the public chat feature was the private chat and needed to be corrected. The other three users were able to find the private chat quickly using the messenger icon metaphor. One of our users thought that having two separate chats was confusing and that we could combine the two, to make all the chatting in one place. Additionally, most of our users did not find it apparent that the list of users in the private chat were ordered by distance.

Some of the improvements a few users suggested we do is to create a way to make a group chat for people within the private chat feature. They also said a notification system and a way to send people emojis or gifs would be beneficial and more appealing. Another 3 users said that sometimes it was not apparent who was speaking in the chat just by name, they thought it might be better if a user profile picture was added to make identifying users a little easier. We also discovered that three out of four of our users preferred to chat using the provided internal chat features rather than an external chat application because it is more convenient.

5.1.6 Notes. The notes feature was clear and self-explanatory. Users were able to open the notes module, create notes and delete them within seconds. They did not need instructions and the layout was understandable. The buttons' meanings and function were clear. One user said it reminded them of sticky notes. One user was not a fan of the UI, suggesting it could have had more aesthetics. Many users had suggestions for additional features for the note's module – like the ability to add persisting names to the notes and the ability to connect a note to a specific artwork.

5.2 Qualitative Analysis

For our qualitative analysis we performed an analytical evaluation. Specifically, we evaluated our project interface and functionality using Nielsen's Heuristics [Nielsen and Molich 1990]; we were able to enumerate specific features of our design supporting each aspect (Appendix A). The usability results of our prototype were positive and reflect well on a successful requirements, design, and prototyping process. Using Javascript, CSS, HTML5 with our Fusality server, We were able to quickly prototype our Metaverse information and interaction designs ideas with X3D and glTF.

6 CONCLUSIONS AND FUTURE WORK

A key to this project was gathering and synthesizing on site observations and user research into our requirements. We then had a short time to design and prototype an interactive multiuser environment to test the different designs. With Fusality, developers can quickly spin up X3D and glTF worlds, leverage multiple parallel Webservices, and also extend the messaging protocol to include shared events for their use case (play, pause, stop, rwd, ffd in-world video).

Future work will include improving of our proximity chat feature, a way to store notes users have written about the artworks, and a space to have a shared viewing experience. Currently the proximity chat feature is only at a lo-fi prototype stage. Additionally, the notes feature temporarily saves the notes in the local storage of the user's browser. A better way of saving and sharing these notes with others could be added. Lastly, the art gallery does not support shared audio experiences. In the future, we would like to see a way for users to form small groups with one another to be able to share an audio or video experience within the virtual gallery.

REFERENCES

Sang Chul Ahn, Ig-Jae Kim, Hyoung-Gon Kim, Yong-Moo Kwon, and Heedong Ko. 2001. Audience interaction for virtual reality theater and its implementation. In *Proceedings of the ACM symposium on Virtual reality software and technology*. 41–45.

Matthew Anderson, Damian Schofield, and Lisa Dethridge. 2013. Ways of viewing and interacting: User experience in a virtual art gallery. *International Journal of Computer Research* 20, 1 (2013), 79.

Kylie Budge. 2013. Virtual studio practices: Visual artists, social media and creativity. *Journal of Science and Technology of the Arts* 5, 1 (2013), 15–23.

Jihyun Kim, Kelly Merrill Jr, Chad Collins, and Hocheol Yang. 2021. Social TV viewing during the COVID-19 lockdown: The mediating role of social presence. *Technology in Society* 67 (2021), 101733.

]Oberstein2021WatchParty title = Lily Oberstein. [n. d.]. ([n. d.]).

Jakob Nielsen and Rolf Molich. 1990. Heuristic evaluation of user interfaces. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. 249–256.

Nicholas F Polys, Benjamin Knapp, Matthew Bock, Christina Lidwin, Dane Webster, Nathan Waggoner, and Ivica Bukvic. 2015. Fusality: an open framework for cross-platform mirror world installations. In *Proceedings of the 20th international conference on 3D web technology*. 171–179.

Bektur Ryskeldiev, Yoichi Ochiai, Michael Cohen, and Jens Herder. 2018. Distributed Metaverse: Creating Decentralized Blockchain-Based Model for Peer-to-Peer Sharing of Virtual Spaces for Mixed Reality Applications. In *Proceedings of the 9th Augmented Human International Conference* (Seoul, Republic of Korea) *(AH '18)*. Association for Computing Machinery, New York, NY, USA, Article 39, 3 pages. https://doi.org/10.1145/3174910.3174952

Scott S Snibbe and Hayes S Raffle. 2009. Social immersive media: pursuing best practices for multi-user interactive camera/projector exhibits. In *Proceedings of the SIGCHI conference on human factors in computing systems*. 1447–1456.

Daniel Volk. 2008. Co-creative game development in a participatory Metaverse. In *Proceedings of the Tenth Anniversary Conference on Participatory Design 2008*. 262–265.

A HEURISTIC EVALUATION RESULTS

Visibility of System Status:

- Users are informed of their location in the virtual world with coordinates at the bottom.
- Users can see the avatars of other players move about the world in real time.
- When users public chat in real time, all avatars in the world can see the messages sent.

No Language Barrier Between the System and the Real World:

- The names and icons we use in our project resemble vocabulary common in the realworld. Terms like avatars, chat

window, world, and map are used to refer to elements on screen, making it easy for users to follow along. Like the "Map" word refers to a map used for navigation within the world.

User Control and Freedom:

- Users can make mistakes in typing the wrong name for their avatar. Fortunately, users can refresh their page to allow the server to go back to the system's previous state allowing them to edit their avatar name.

Consistency and Standards:

- This product uses common practices used in web applications. For example, the map icon has the image of a folding map. It also uses pins as checkpoints similar to what is used on Google Maps.
- The notes feature is like the notes app found in various systems.
- The "Messenger" icon for private chatting.

Error Prevention:

- When the user enters a world, a pop us appears asking them their name. If they make a mistake by not assigning a name, the system automatically gives them a name like "User 1" so that the user is initialized and not null.

Recognition, Not Recall:

- The user can easily see all the options available to them on the screen in a distinctive manner, so they do not have to remember anything.
- They are easily recognizable as well like our avatar option, online users list, map button, chat window, and message button.

Flexibility and Efficiency of Use:

- The map contains checkpoints users can access to get to a specific location quickly with the press of a single button.
- Each artwork also contains the same teleportation functionality allowing them to change viewpoints and see the artwork instantly.

Aesthetic and Minimalist Design:

- Our private chat button and map button are placed in the bottom right corner in our "tool pane". This keeps our design simple.
- The public chat feature is a sidebar allowing users to toggle open and close the chat box and type their messages easily as well as saving space.

Help and Documentation:

- A README form can be found on our GitHub page [https://github.com/SamyCoder/theArtMetaverse] detailing the features of our product as well as which files they can be found in.