Evaluating Collaborative Cues for Remote Affinity Diagramming Tasks in Augmented Reality

Nathaniel R. Llorens

Thesis submitted to the Faculty of the
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of

Master of Science
in
Computer Science and Applications

Sang W. Lee, Co-chair
Douglas A. Bowman, Co-chair
Wallace S. Lages

August 12, 2021
Blacksburg, Virginia

Keywords: augmented reality, brainstorming, affinity diagramming, collaborative cues

Evaluating Collaborative Cues for Remote Affinity Diagramming Tasks in Augmented Reality

Nathaniel R. Llorens

(ABSTRACT)

This thesis documents the design and implementation of an augmented reality (AR) application that could be extended to support group brainstorming tasks remotely. Additionally, it chronicles our investigation into the helpfulness of traditional collaborative cues in this novel application of augmented reality. We implemented IdeaSpace, an interactive application that emulates an affinity diagramming environment on an AR headset. In our application, users can organize and manipulate virtual sticky notes around a central virtual board. We performed a user study, with each session requiring users to perform an affinity diagramming clustering task with and without common collaborative cues. Our results indicate that the presence or absence of cues has little effect on this task, or that other factors played a larger role than cue condition, such as learning effects. Our results also show that our application's usability could be improved. We conclude this document with a discussion of our results and the design implications that may arise from them.

Evaluating Collaborative Cues for Remote Affinity Diagramming Tasks in Augmented Reality

Nathaniel R. Llorens

(GENERAL AUDIENCE ABSTRACT)

Our project was aimed at creating an app for modern augmented reality headsets that could help people perform group brainstorming sessions remotely from each other. We were also interested in finding out the benefits or downsides of some of the design decisions that recent research in remote augmented reality recommends, such as lines showing where a user is focusing and visualizations for a user's head and hands. In our app, which we dubbed IdeaSpace, users were faced with a virtual corkboard and a number of virtual sticky notes, similar to what they might expect in a traditional brainstorming session. We ran three-person study sessions comparing design techniques recommended by literature to an absence of such techniques and did not find they helped much in our task. We also found that our application was not as usable as we had hoped and could be improved in future iterations. We conclude our paper discussing what our results might mean and what can be learned for the future.

Acknowledgments

Dr. Lee, Dr. Bowman, and Dr. Lages for their invaluable guidance.

Tom Phan for his helpful advice and work on this project.

My mom and dad for their encouragement and support.

My brother for spending time with me when I was stressed.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

In this chapter, we introduce important concepts and provide the motivation for our project. Additionally, we set out our aims and the research questions we answered.

1.1 Brainstorming and Affinity Diagrams

Before making important design decisions or choosing a solution to a problem, it is common for a design team to hold a brainstorming session. Group brainstorming is a group creativity process that involves team members cooperating to record their thought processes as they attempt to discover potential solutions to their problem [20]. Inherently, brainstorming is a collaborative process, and often team members will each post their thoughts on small notes and append them to a common board [20]. As a potential solution takes shape, the notes might be reorganized and restructured.

Affinity diagramming [13] is a process that can be similar to a group brainstorming session. The process involves trying to find patterns in unstructured data and organizing the data into clusters [22]. Often it is a process that can be performed on the output of a brainstorming session, where many ideas have been generated and now need to be structured by affinity, or similarity.

Traditional brainstorming has its limitations. Typically, brainstorming and affinity diagramming are performed on physical sheets of paper, such as small sticky notes, and placed or grouped on a board that acts as the group sharing space. Because of its physical nature, ideas cannot easily be modified or placed into multiple different groups. Additionally, the output of brainstorming sessions cannot easily be digitized, as this would involve intensive typing or taking digital pictures that do not include important textual data, limiting the abilities of users to search or iterate over the output. Digital tools and methods, such as collaborative augmented reality, could be used to transcend these limitations and improve the processes.

1.2 Collaborative Augmented Reality

Augmented reality (AR) is a medium that has gained increased prominence in recent years due to advances in commodity technology and lower prices [6]. It provides a higher level

Figure 1.1: Microsoft HoloLens AR headset (credit: https://www.roadtovr.com/microsoft-hololens-augemtend-reality-virtual-reality-headset-microsoft-holographic-windows-10/)

of immersion than the two-dimensional screen of a computer or phone, especially when used with a headset, but also does not completely remove a person from their physical surroundings. AR is similar to virtual reality, but it involves starting first with the physical world and then superimposing computer generated graphics onto it [17]. Typically, AR also involves some interaction between the virtual and physical world, even if the interaction is as small as anchoring virtual objects to a physical location. AR can be a powerful tool, especially for overlaying additional information from the world around the user that the user would not have been able to see otherwise — e.g., the 3D model of an underground mine that an app allows a user to view from the surface. In Figure 1.1, we show the Microsoft HoloLens, an AR headset that debuted in 2016 with a price of $3000 for the Dev kit and $5000 for commercial [12].

Despite being a powerful medium, collaborative AR comes with its own set of challenges. Immersive virtual environments can often require additional cues that would not normally be considered or taken for granted in the real world. Kim et al. [14] showed the importance of pointers and annotations in creating an experience that feels truly collaborative, where each user feels as if they were sharing an environment together. Liao & Wang [16] demonstrated how the additional visibility of hand gestures in a collaborative task can improve communication between parties. Indeed, much of AR research is centered around constructing affordances to make up for the lack of visual body cues or other kinds of cues taken for granted in the real world [7, 9, 19]. Despite these challenges, we believe AR can be used to augment group brainstorming.

1.3 Aims and Objectives

The aims of this thesis are two-fold. The first is to evaluate the potential of AR for effective brainstorming or affinity diagramming with remote users. The second is to observe the performance of commonly used collaborative cues in a brainstorming task with remote users.

1.3.1 Aim 1

To accomplish the the first aim, the work can be split into a number of specific objectives. The first objective is to create a virtual environment for AR headsets in which a user can manipulate virtual sticky notes around a board. Secondly, the environment needs to be networked so remote users can interact in the remotely shared environment. Lastly, collaborative cues need to be implemented so the remote users will be able to visualize what each of them are doing and what they are manipulating.

1.3.2 Aim 2

The second aim involves altering the app from the first aim. The collaborative cues need to be toggle-able so the investigator can enable or disable them during a study session. Additionally, prefabricated notes need to be generated so that the study participants will have notes to cluster. Finally, a new logging system needs to be added to the app that will log interesting interactions to a text file.

1.4 Research Questions

Along with aims, we also had two research questions that we sought to answer in our user study.

RQ1: To what extent can AR-based affinity diagramming be usable?

RQ2: How do awareness cues or affordances for remote collaborative AR effect performance and perceived quality of collaboration in a collaborative affinity diagramming task with multiple users?

We consider the answers to these questions in the Discussion chapter.

1.5 Contributions

In this thesis, we created a tool for brainstorming tasks, which we dubbed IdeaSpace. The main contribution of our tool is the exploration of the potential of AR for remote affinity diagramming. Additionally, we observed the inclusion or exclusion of common collaborative cues in our tool to see how they fared, and found these cues had little impact on the clustering tasks our participants completed, likely due to learning effects which we address in the Discussion chapter.

Chapter 2

Review of Literature

In this chapter, we summarize past works relating to our project.

2.1 Remote Collaboration

In recent years, AR and virtual reality (VR) technology has become increasingly accessible to consumers and researchers thanks to dropping prices and improvements in usability [6]. As a result, collaborative AR and VR have seen increased popularity in the last ten years of research.

Although collaborative AR has become popular more recently, remote collaboration is a mature field where researchers seek to find the next best ways to design applications and interfaces to take advantage of emerging technologies and methodologies. In this section, we explore a number of works published regarding remote collaboration.

Higuchi et al. [10] created a digital whiteboard that was meant to simulate a real whiteboard for remote collaboration. There were two prototypes, each consisting of a large touch screen with a Kinect camera on one end or the other depending on whether it was the right or the left board. The authors later tested their prototypes to see how well users could interact from one board to the other using a teaching and a game task.

Tang et al. [23] attempted to design a tabletop interface that could support three-way remote collaboration. After creating a program that ran on touchscreen tabletops, they performed two user studies. The first user study was intended to determine if different task types might favor different configurations of users in the workspace. They found that the same-side configuration, where all users were represented as being on the same side of the table, was best for text-based work and the around-the-table configuration, where the users were spread in various positions around the table, was best for non-text-based work. Their second study involved comparing media channel configurations for which was optimal in test tasks. They found that the cues they offered worked reasonably well.

Liao & Wang [16] sought to understand the importance of hand gestures in computer mediated brainstorming. They constructed a user study with three conditions and compared them. Their hypothesis that users would be more creative in media where they could see each other was not supported by their results. However, they did find that users used hand gestures more in these kinds of media than in audio only. It also seemed that gestures were helpful for creativity. They used their findings to recommend hand gestures be incorporated into computer mediated brainstorming programs in order to benefit users.

Gauglitz et al. [8] created a software application that could support collaboration between a local worker and a remote expert. They were interested in discovering whether enabling the remote expert to create in-world annotations that the local user could see in the desired position would aid in a simulated real-world task. They designed a study mimicking a real-world scenario that a remote expert would guide a local user through and compared their world-stabilized annotation technique to a static annotation technique and a normal video conferencing style. Users preferred their technique, but it was not significantly better than the static annotation technique.

Scott & Inkpen [21] were interested in how people divide up tabletop spaces and whether this knowledge could be used to improve tabletop digital interfaces. To investigate how users segment tabletop spaces they held an observational study and identified three separate area types: personal, group, and storage. Using these observations, they held a user study where they segmented the tabletop task space and recorded how often actions occurred in the various segments. This study helped them get a better idea of the area each of their identified categories covered. Finally, they condensed their findings down into design guidelines of future tabletop interfaces.

Widjaja & Takahashi [24] chose to design a distributed interface to support affinity diagramming between remote users. They used a two-screen configuration, where one screen would display the authoring space and the other would display the public space that all users could see. They color coded authored ideas by user to help improve productivity. After designing their interface, the authors held user study sessions that involved discussions of topics and the use of their interface to present ideas, perform affinity diagramming, or other tasks. They concluded from their experiments that the users were able to come up with more ideas using their digital means than physical means.

Similar to Gauglitz et al. [8], Kim et al. [14] noted a number of problems with video conferencing systems and sought to improve two factors they identified as "sense of being together" and "connectedness". They designed a prototype remote collaboration system, with one user being considered the local user and the other considered the remote user. The local user would be a user wearing a head-mounted device (HMD) or holding a handheld device (HHD), and the remote user would be a user at a desktop computer, able to see the video from the local user and make annotations or point to things. They conducted a user study where they compared a number of conditions and found most users preferred the pointer condition most, with the baseline coming in last.

Brennan et al. [3] were interested in improving visual search tasks with pairs of people. They implemented a system for shared gaze that displays a gaze cursor for each user of the user's eye position, obtained from eyetrackers. Additionally, their system supported audio communication between the two searchers. The task in their study was for the pairs to collaboratively search for a letter O in a group of Qs. The study was also configured with various communication conditions to try to gauge the effectiveness of each. They found that users in the no-communication baseline (with no shared gaze or audio) showed far more search misses than the other conditions. They also found users seemed to use various strategies for spatial division of labor in the communication conditions. Interestingly, they found that adding speech actually hurt rather than helped performance in the task. The authors ultimately concluded that collaborative search is faster than solitary search and that the shared gaze was effective enough to mediate collaboration.

Much like Brennan et al. [3], D'Angelo & Gergle [5] investigated a system of shared gaze in order to improve collaboration in a paired task. They developed a study with three settings: co-located, remote with shared gaze, and remote without shared gaze. The task they gave their study participants was to assemble a puzzle, where users on computers could move puzzle pieces by selecting two pieces at once. Shared gaze was displayed as a cursor captured by eyetrackers with a smoothing function.

D'Angelo & Begel [4] devised a system to improve pair programming, a common practice at technology companies, using shared gaze visualization. However, for this project they used a processed eyetracking method in order to help slow down the visual cursor to mitigate biases in the technology. Once they had implemented their method, they performed a study with pairs of software developers and had them perform refactoring tasks with their visualization method enabled and disabled. Following the study, the authors analyzed recorded screen capture video. They found that users of their gaze visualization method had more gaze overlap which indicated improved collaboration. Completion time seemed to stay the same between the two conditions. They concluded their work by pointing out that their design improved collaboration and was unobtrusive. They found participants used the gaze cursor to refer attention to specific pieces. They also found, however, that the shared gaze could be detrimental, especially when it was somewhat misaligned, causing the eye cursor to be displayed over the wrong puzzle piece. The main contribution of the authors' work was their recognition that users of their shared gaze were able to collaborate and coordinate in similar ways too the co-located users.

2.2 Collaborative AR

AR has been a topic of research for far longer than it was a practical reality. Novelists envisioned a world where people could be surrounded by AR interfaces that helped them in their everyday lives [17]. As AR technology matures and researchers explore it further, it is important to have knowledge of what came before and what might await the field in the future. In this section, we explore two works that provide brief histories of AR and give predictions of its future.

Lukosch et al. [17] set out to document past visions of AR, the state of AR at the time, and the work in collaborative AR that still needed to be done in 2015. They described myriad scenarios and tasks where AR had been experimented with and the many ways in which it could improve collaborative tasks that otherwise would be limited by the physical world or commonly used digital technology. They emphasized the need for researchers to further explore the realm of remote AR collaboration, discovering the ways in which the presence and awareness of remote users could be improved. They also pointed out that research into utilizing senses beyond simply aural and visual was rare and lacking.

Similar to Lukosch et al., Ens et al. [6] performed a survey of collaborative mixed reality in 2018. The authors were seeking to provide an overview of collaborative mixed reality up until the time of writing, explain what regions and topics had not been explored yet, and give predictions for what they felt the future of collaborative MR would look like. Reflecting on thirty years of research, they found that much modern MR research is clustered under two scenarios: remote expert [9] and shared workspace. The remote expert scenario generally consists of a local user who is a novice receiving guidance from an expert located elsewhere. Often the remote expert receives a view of where the novice is located and what they are looking at to better assist in guiding the novice in their task. Shared workspace scenarios have a different dynamic, where multiple users (typically only two in user studies) are present in the same virtual workspace. These scenarios tend to have a remote focus, with the two users not necessarily being located in the same physical space, but still experiencing the same virtual space.

Ens et al. [6] believed that going forward there would be more of a diversification of roles from the remote expert and shared workspace scenarios as MR became more popular. They also predicted that MR would become overall more flexible, allowing users to shift between different modes at will, such as synchronous to asynchronous and vice versa. Additionally, they predicted a convergence between the digital and physical worlds.

Men & Bryan-Kinns [18] created a shared virtual environment that they dubbed LeMo. Within LeMo, two users could collaboratively create music using VR headsets. LeMo's environment consisted of a music interface, the avatars of each user, and the virtual space itself. The authors developed and performed a study where they expected territoriality to emerge between the users and thought personal spaces would help the users' collaboration be more efficient. Their hypotheses were supported by their experiment. Their given takeaway from their work was that personal spaces are important for collaboration and that visible personal space is preferred.

2.3 Cues in Collaborative AR

Much research has focused on studying the role of cues in collaborative AR. Part of the challenge of multi-user AR and VR is making up for the loss of typical cues that humans rely on in the real world, such as body or facial cues. Researchers experiment with compensating for this lack by adding their own cues, such as a ray indicating the gaze direction of a remote user [19] or a visualization of the remote user's hand state.

Li et al. [15] identified a problem in collaborative AR: namely, the lack of correct occlusion cues. In a co-located setting, where users are together both physically and in the virtual world, it becomes important for each user to accurately see when the other user's hand should occlude a virtual object. However, computing the position of a user's hand in real time is expensive, resulting in often inaccurate positioning. The authors were interested in studying the effects of such inaccurate occlusion cues, so they performed an experiment with different seating arrangements in which one user references numbered cubes (physical or virtual) and the other observes. They found that incorrect occlusion had an impact on accuracy and task completion time and that the participants attempted to compensate for the bad occlusion using different strategies. They recommended a few ways designers could compensate for incorrect occlusion, such as by dispersing content in front of viewers rather than along their view lines.

Jing et al. [11] investigated the effects of gaze visualization in co-located AR. Their gaze visualization system featured four different configurations of gaze cues: non-collaborative (the baseline), cursor donut (where both gaze locations were indicated with cursor donuts), laser eye (where remote user gaze location was indicated with a cursor donut and laser or ray), and trail path (where remote user gaze location was indicated with a cursor donut and a path that the gaze has traveled). Additionally, the cursors and paths changed colors depending upon the amount of ”dwell time” upon an object. They performed a user study with three different task types and participants working as pairs. They found that users preferred the laser eye configuration, and also that the laser eye configuration was better for understanding partner’s focus and helping users react to their partners. Laser eye also had the lowest average completion time. Users expressed a desire to be able to toggle visibility of their or their partner’s cues, but the authors still concluded the cues were not very distracting and helped users in their tasks.

Oda and Feiner [7] attempt to address the problems and limitations of pointing mechanisms in collaborative AR, specifically when pointing at physical objects. They call their approach GARDEN and it involves the ‘indicator’ first using a sphere to select an area in the physical world, and then a reconstruction of that area being generated close to them so they can point at specifically what they wish to point out to the ‘recipient’. The recipient sees a reconstruction of the indicator’s hand or fingers with their copy of the selected area. Their method uses a Kinect depth camera to model the physical space and track the indicator’s hands. They compared their technique to a number of other techniques and found that although it was a bit slower, it seemed to have significantly better accuracy.

Piumsomboon et al. [19] compared combinations of cues, such as Field of View (FoV) frustum, FoV frustum and eye-gaze ray, and FoV frustum and head-gaze ray, against a virtual representation of a user’s head and hands and found that the eye-gaze and head-gaze configurations outperformed the head-and-hands configuration. Their findings perhaps imply that additional cues are needed beyond just head and hand position, likely due to the low FoVs of contemporary AR and VR headworn displays (HWD).

Kunal et al. [9] put together a head-mounted AR system that also incorporated eye tracking. In their experiment, a remote user watching a computer monitor could see the task space via cameras on the head-mounted display, as well as the point where the task user’s eye was focused. From the results of their experiment they found that their system scored higher on their Likert scale survey than other baselines. Additionally, they found that tasks were completed faster with both eye tracking conditions. Users seemed to prefer the condition with both their pointer and eye tracking cues, and this was borne out in their results as well. The authors conclude that these kinds of cues can provide an improvement in such collaborative tasks.

Chapter 3

Design Goals

The first major portion of this thesis is aimed at producing an AR application running on contemporary hardware that utilizes collaborative cues to facilitate affinity diagramming tasks. In this chapter, we identify specific goals for our design to fulfill.

3.1 Emulate Physical AD Environment

Past research has shown that providing familiar interfaces to users in new forms of technology can show large benefits [16]. Such interfaces can leverage existing knowledge or concepts in the users, cutting down on the overhead of learning a new system. Since we are creating an application for performing affinity diagramming tasks, one of our goals is to include similar structures to such a task in a physical setting.

One conventional affinity diagramming structure is the clustering space. Depending on the affinity diagramming variant, the form this space takes may vary, but at its simplest it can take the form of a board, such as a whiteboard, walls, or corkboard.

Another structure that any affinity diagramming user would be familiar with is the sticky notes where users write their ideas. In computer-based affinity diagramming software (such as Miro, see Figure 3.2) the authored notes might take the form of a square or rectangle with text or other data inside it, whereas in a physical setting it is likely to be sheets of paper or

Figure 3.1: An example of affinity diagramming (credit: https://ivanvasilev.com/blog/affinity-diagram/2020/)

sticky notes (as in Figure 3.1). These forms intend to 1) enable the users to express their ideas with enough granularity and 2) enable them to easily move and cluster the ideas.

In the Software Design section, we describe how we included these structures in the application. For the board, we implemented a virtual corkboard that supported snapping of ideas onto it. The ideas we represented as virtual objects resembling sticky notes that could be appended to the virtual corkboard when brought close enough to the surface (see Figure 4.2c). A brainstorming user using our application for the first time would have these two known concepts, decreasing the time needed to learn the new interface.

3.2 Augment Natural Interactions

Similar to our reasoning in the previous section, we decided it would be useful to augment familiar input mechanisms for users to interact with the virtual affinity diagramming setting.

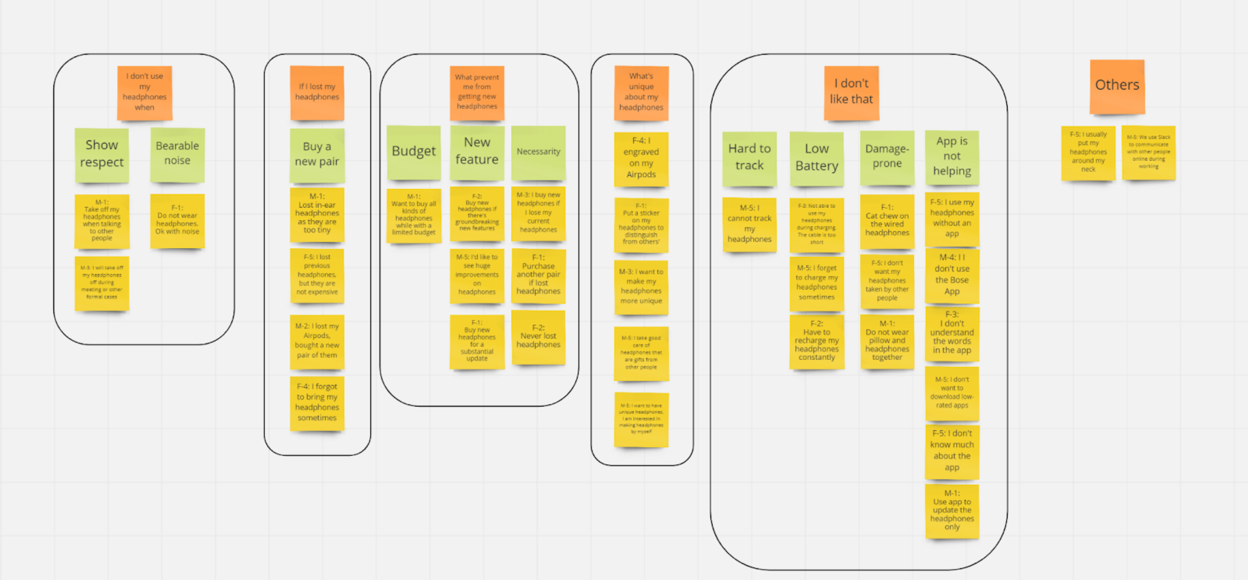


Figure 3.2: Miro online whiteboard (credit: https://blog.prototypr.io/how-i-redesign-the-sony-headphones-a70b485a9b08)

Because our focus was on the collaborative clustering aspect of affinity diagramming, we identified a few important interactions for our application to support. The first and most important was the ability to manipulate ideas and move notes around the environment, preferably in a way similar to physical affinity diagramming using a user's hands. As we discuss in Software Design, we augmented this interaction by enabling users to select a distant virtual sticky note with their gaze and then to move it with their bare hands. We chose this method so users would not have to move to a note in order to manipulate it. Another interaction was posting authored ideas to the central clustering space; in other words, enabling a user not just to move an idea but also to arrange it within the clustering space. By enabling users to move virtual sticky notes and automatically snapping them onto the virtual corkboard when they have been moved close enough, we supported this second interaction.

3.3 Support Non-Verbal Communication in Remote Settings

Our last design goal was to make sure our application enabled users to perform a collaborative affinity diagramming task even if all users are remote. In addition to synchronizing world state for each of the users, it also provides collaborative cues that enable users to work together rather than each working separately.

As we discuss in the following chapter, we implement each of these cues in the cue condition by providing a line from each user's head position to the virtual object of their focus and a sphere for each user's head and hand positions. Each cue we implemented is color coded to match the user it belongs to.

Chapter 4

Software Design

In this chapter, we document how we put the design goals from the previous chapter into practice. We also provide an overview of the software architecture of our application.

4.1 Development Environment

IdeaSpace was developed on a Windows 10 PC using the Unity 2019.4 game engine. We also use the Photon Unity Network for peer-to-peer cloud networking and Microsoft's Mixed Reality Toolkit (MRTK) 2.5.1. C# script code was written and edited using Microsoft Visual Studio 2019. The application was tested and deployed to three first-generation Microsoft HoloLens devices.

4.2 IdeaSpace

Our shared AR environment is designed after a common brainstorming scenario. Often in a physical brainstorming session, team members document small simple ideas and display them for other team members to see. Additionally, ideas might clustered together to indicate close notes are similar in ideas. To facilitate these features, we decided to place a two-dimensional display board in the scene and to allow users to arrange virtual objects on it however they desired. To represent the ideas, we introduced virtual sticky notes as movable yellow squares

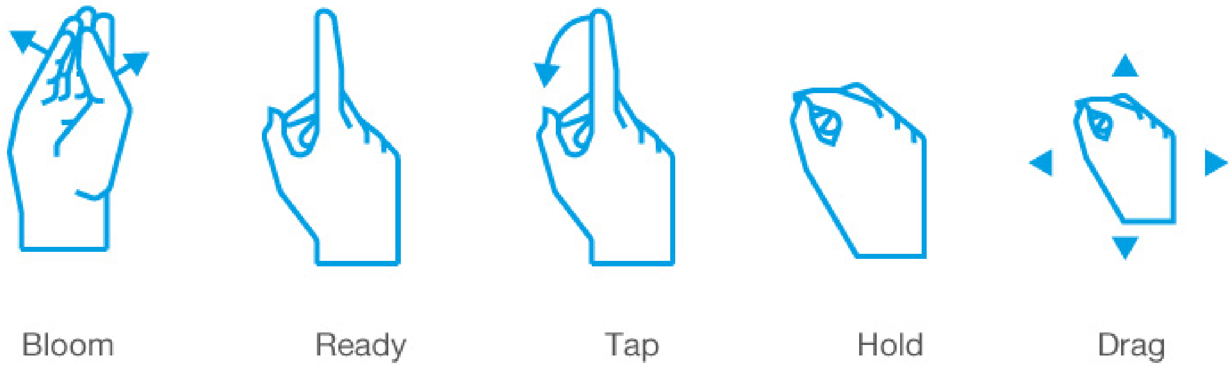
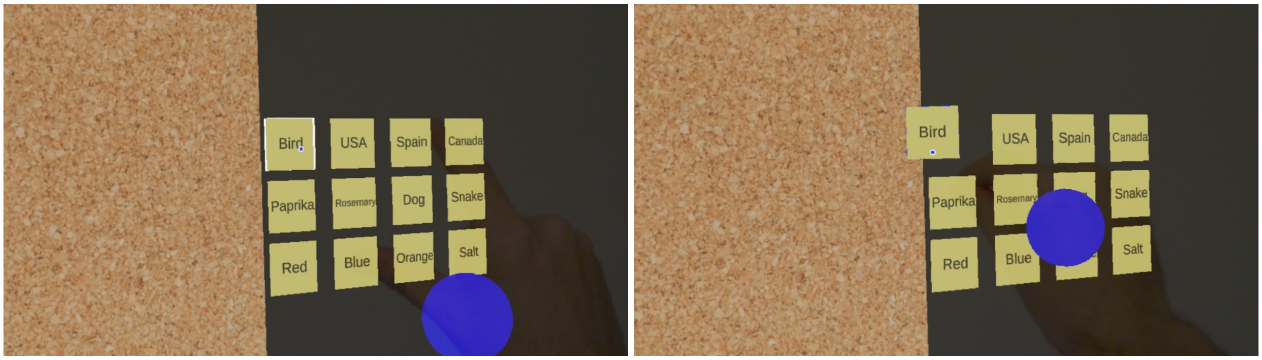


Figure 4.1: HoloLens gestures (credit: https://msd-makerspaces.gitbook.io/next-lab/augmented-reality/resources/platforms-hardware/microsoft-hololens)

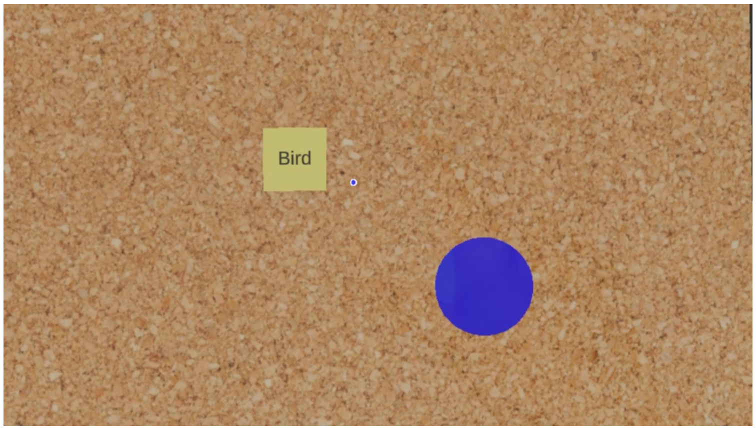
with text on them.

It was important that our application augmented similar manipulations to those used in the real world. Thus, IdeaSpace includes two methods for users to interact with virtual sticky notes. The first involves the user performing the HoloLens Hold gesture (see Figure 4.1) with their cursor positioned on a note, locking the note to the center of the user's screen and allowing them to move the note at a fixed distance by simply turning their head. A simple HoloLens Tap gesture stops the note's movement when the user is satisfied with the placement, similar to the drag-and-drop gesture commonly used with a mouse on a personal computer. The second method is more hands-on and is ultimately the method that was used in our study. While having the HoloLens cursor positioned over a note, the user performs a HoloLens Drag gesture, allowing them to move the note by moving their hand. To clarify, the note does not follow their hand, but rather matches the movement of their hand, scaled by a multiplier of 3.0, similar to Bowman et al.'s Go-go method [2] but linear. Initial versions used a one-to-one mapping, but we found such a low sensitivity made moving notes across the board difficult, so we experimented with values until settling on a value of 3.0. As soon



(a) User puts stationary note in focus. Notice the white highlighting around the note

(b) User starts Drag gesture and commences movement of note. Notice the blue highlighting around the note, signalling blue local user is moving the note



(c) User stops Drag gesture and ends note movement. Notice that there is now no highlighting on the note

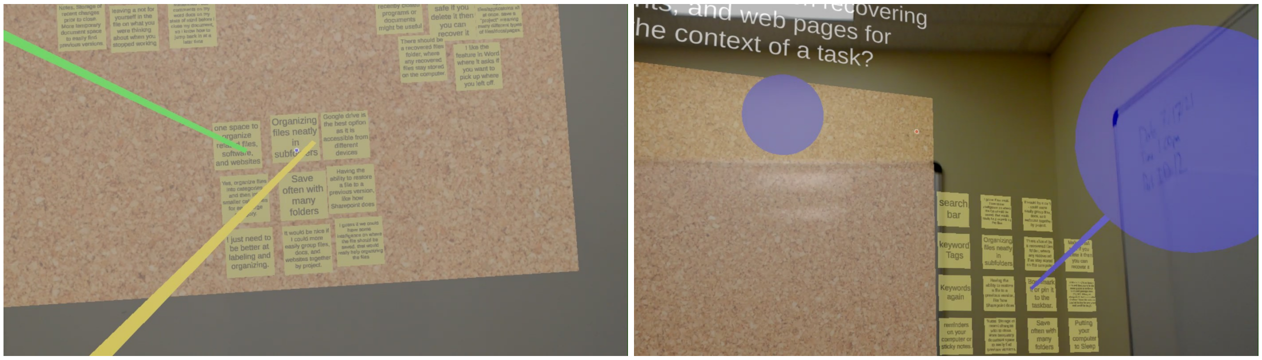
Figure 4.2: Note movement process from start to finish. The blue sphere is the hand visualization for the local user.

as the user stops the Drag gesture, the note stops moving. This process is shown in Figure 4.2.

One additional feature that our application supports, but that was removed in the version used in our user study (because it was not within the study scope), is the ability to create a blank virtual sticky note positioned near the virtual board. The user place their cursor anywhere overlapping the board, and when they performed a HoloLens Hold gesture, the blank note would appear at the location of their gaze.

4.3 Networking and Synchronization

In order to support multiple devices used by multiple users, any collaborative AR app requires data communication to keep device states in sync. Because it was well-supported and to avoid the need to write a server version of the app, we settled on the Photon Unity Network (PUN) third-party library, a peer-to-peer cloud-based networking solution, for synchronization. With a few simple lines of code, peers could automatically connect to a PUN room unique to our app. Because of its peer-to-peer nature and concept of object ownership, synchronizing state through PUN had to be initiated from whichever peer owned a specific virtual object. To accomplish this, virtual objects whose state needed to be shared were equipped with a PhotonView component, a PUN helper class that exposes networked object properties such as ownership and automatically synchronizes serialized data [1]. This component is used to observe simple components like PhotonTransformViews, a PUN helper class that observes the position and rotation of a virtual object, to automatically synchronize object position and rotation [1]. PhotonViews could also be used to invoke Remote Procedure Calls (RPCs). RPCs allow remote peers to call methods on other peers and enable sending parameters through the Photon Cloud for those methods. This offered us a more manual approach to synchronizing state than a PhotonTransformView and enabled us to ensure that things like text, color, and certain flags of objects were correctly synchronized on other peers when a change was triggered in one peer.



(a) Gaze ray examples (b) Gaze ray extending from head visual

Figure 4.3: Two examples of gaze rays. 4.3a illustrates two rays in the user view. 4.3b shows a gaze ray extending from a user's head visual. The small sphere is the visualization of user's hand, and the large sphere is a visualization for their head.

4.4 Collaborative Cues

In a co-located environment, where all AR users are located in the same room, visual cues for awareness are not always necessary, but when the same app supports remote participation, users require information from the app about what others are doing or experiencing. In much of the research on collaborative AR, a similar set of cues are used [6, 9, 19]. Perhaps the most common is the gaze ray, which is a simple ray that extends from a user's position in the virtual environment out to the point the user is focused upon (see Figure 4.3). This approach does not account for the fact that a user's eyes may be focused somewhere other than the center of the AR device's screens, but it seems to function as an adequate estimate of focus [19]. Another commonly used visual cue is head and hand visuals. In an app where users are able to rotate their heads and move around a virtual environment, it becomes important to visualize the location and orientation of the head and sometimes the hands to help make up for the lack of physical body cues [19].

In our environment we focused on implementing the common cues we found in literature, namely gaze rays and head-and-hand visuals [19]. Gaze rays were fairly trivial, only requiring

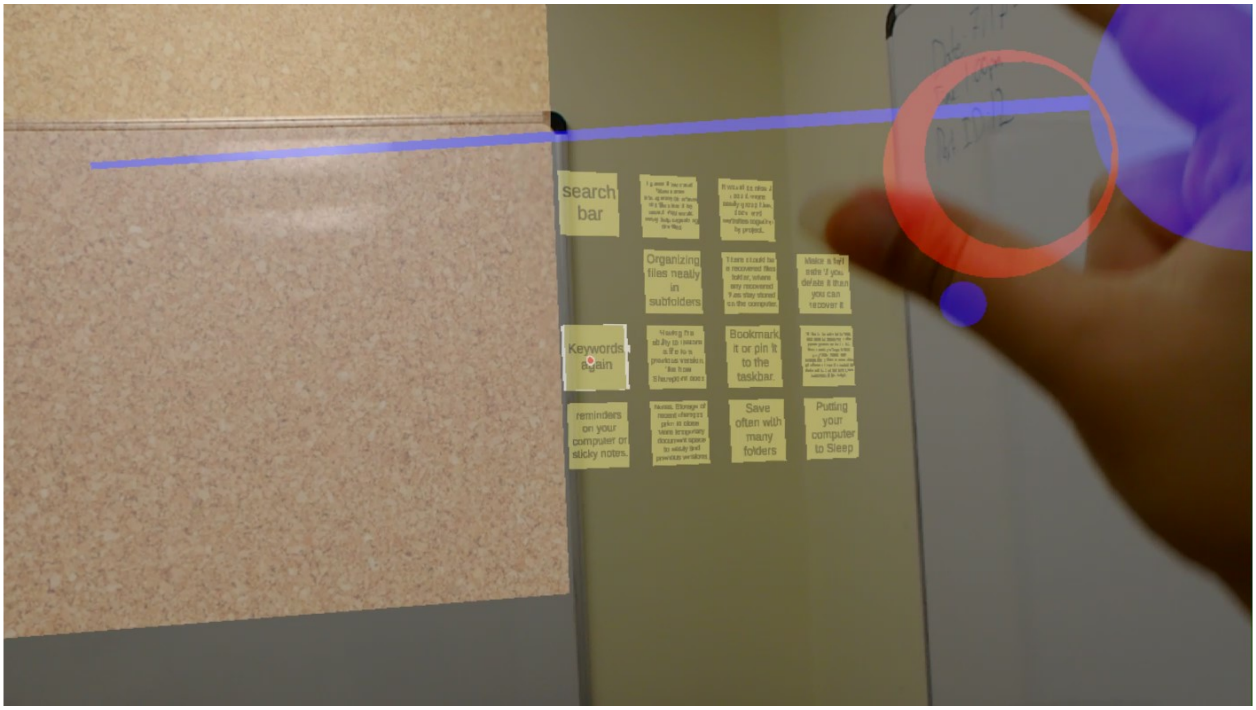


Figure 4.4: Note highlights white when user cursor hovers over it. The blue spheres are hand and head visuals for another user, and the red sphere is the hand visual for the local user.

a line from the user's head position out to whatever object the user was focused upon. The end point of the line was found using a ray-cast on the object to prevent lines from penetrating objects and potentially causing ambiguity. We then ensured that the gaze ray of a user only displayed to other users, but not in their own view. Head-and-hand-only visuals offered more of a challenge. The head visual was as simple as keeping a sphere at the reported position of the HoloLens within the world, but the hand visual required accessing the underlying controller that the Unity Engine uses to represent articulating hands. Once the application was able to listen to movement events fired when the HoloLens detected hands in its view, matching the hand location with a sphere became trivial. To try to match the approximate sizes of a user's head and hand, we applied scalings of 1.0 and 0.05 to the spheres, respectively. Figure 4.3b illustrates how these cues appeared to the users.

Additionally, we chose to include a couple of UI features that would hopefully provide useful

(a) Remote green user moves note

(b) Local red user moves note

Figure 4.5: Examples of notes highlighting by user

information to users. One such feature is the color coding of each of the cues so users can have an idea of whose ray or head they are viewing, as shown in Figure 4.3a. Another simple UI feature is colored outlining of virtual sticky notes. There are two main outlining states: white borders when a user's cursor is over a note to help them understand which note will be affected by their manipulation, as shown in Figure 4.4, and color-coded borders when a note is being manipulated by other users so each user understands who is moving a note (see Figure 4.5).

Chapter 5

User Study

In this chapter, we discuss the user study that we conducted as well as the pilot study that led up to it and informed some of our design decisions.

5.1 Pilot Study

We held various testing and demo sessions before any true pilot study. In the early sessions, the app was in need of serious user interface changes, and we implemented them incrementally until we felt it was ready for a pilot study with participants who were not familiar with the tool.

5.1.1 Participants

We recruited two participants (one age 18-24, the other age 25-34, both male) for the single pilot study session, with the third being an investigator. The participants were recruited from the Virginia Tech graduate student mailing list. Neither participant described themselves as at all familiar with the Microsoft HoloLens.

5.1.2 Study Procedures

Due to the Covid-19 pandemic during our study, all devices were wiped and disinfected before arrival of participants. Additionally, a Zoom call was configured between three different rooms to allow for verbal communication, as well as shared video, for our “remote” study. When the participants arrived, we handed them hard copies of the same consent forms that they had been sent via email days prior. The participants were given as long as needed to read through the documents again before signing. After consenting to be part of the study, the participants were instructed on how to equip the HoloLens. They were also given a cursory tutorial on the hand gestures they would need to use for the HoloLens to recognize their inputs. Each participant was then moved to a separate room, where they were still able to communicate via the Zoom call. They were instructed to open the application, putting each user in the virtual environment. They were given 5 minutes to complete a training task of clustering notes with simple text on them, such as “Dog” or “Purple”. All collaborative cues were visible in the training task. The participants were not told of the categories, but told to guess them, and for this task the categories were plain from the data. Once all the virtual sticky notes were clustered based on categories on the virtual board, the investigator spoke a voice command to reset the environment with a new dataset but the same cue configuration. This was the first experiment task, and the participants were given 20 minutes to cluster the notes in the same way as the training task, but with different categories that they came up with. Finally, once this was completed, the last task was introduced, but for this task none of the collaborative cues were visible. The dataset was similar to the one used for the first experiment task, but not the same to avoid biasing. They were given another 20 minutes to complete the task, and once they reported it was completed, they were allowed to unequip the HoloLens devices and return to the main room, where they were asked open-ended interview questions, whose responses were recorded.

5.1.3 Findings

We found that the users sometimes had trouble moving notes, but this seemed to be a limitation of the HoloLens having difficulty recognizing hand gestures and not a software problem. Overall, the user experience seemed acceptable to continue on to a user study.

Following the pilot study, we decided to make some changes to the app and to the study structure. One change to the study structure was including a second training task between the experiment tasks so that participants could get used to the cue-less condition before needing to use it in earnest. A change to the app that we made following the pilot study was to include a visual for hand position for each of the users in the cue-full condition. Additionally, we added a logger for interaction traces such as when a user looks at a note or when a user moves a note. The output file for this logger could be collected from the devices following the study.

5.2 User Study

Following our pilot study and implementation of the changes necessitated by it, we felt we were ready to test our application in earnest. In this section we describe the makeup and procedure of our user study.

5.2.1 Participants

We recruited twelve (N=12) participants and gathered some demographic information about them via our pre-experiment survey. Eleven were ages 25-35 years old and one was aged 18-24 years old. Seven of the participants were male and five female. Many of the participants claimed to be slightly familiar with AR headsets, but only four reported that they were

Session Number	First Condition	First Dataset	Second Condition	Second Dataset
1	Cue	1	Cueless	2
2	Cue	2	Cueless	1
3	Cueless	2	Cue	1
4	Cueless	1	Cue	2

Table 5.1: Cue conditions and datasets for each user study session. Note that each session had two conditions, hence the first and second conditions and datasets.

familiar with the Microsoft HoloLens. None but one of the participants had ever participated in an affinity diagramming session.

We held four user study sessions with three participants taking part in each. As we will explain further in the following section, each session was split in half, with one half being held configured in one condition and the second half being held configured in another. In order to prevent bias, we made sure to have a session in each valid permutation of cue condition and dataset. Table 5.1 explains the order of condition-dataset pairings used for each session.

5.2.2 Experimental Conditions

IdeaSpace was configured in two different cue conditions during our user study: cue and cueless.

Cue Condition

Our cue condition provided users with common collaborative cues. Gaze rays, head and hand visuals, and note highlighting were all present in this condition.

Cueless Condition

The cueless condition did not provide users with any of the above cues. Gaze rays, head and hand visuals, and note highlights were all disabled. Users were still able to see the HoloLens cursor, however.

5.2.3 Datasets

We used four total different datasets during our experiment: two during the training portions, and two during the task portions.

Training

The training datasets we used in the study were intended to be simple to cluster so users could instead focus on the task and practicing note movement. As a result, we made the datasets fit neatly into four different categories: colors, spices, countries, and animals. Some examples include “Red”, “Salt”, “Canada”, “USA”, “Rosemary”.

Task

The task datasets were deliberately more ambiguous than the training datasets to promote collaboration and help to make the tasks harder. They consisted of open-ended survey responses from previous work (used with permission) that answered one of the following questions:

1. Do you have any ideas (e.g., a feature that you would like in a web browser) in effectively organizing and curating web pages?

Example Data:

- I do like Safari's "frequently visited" webpages that they collect for you as kind of suggested bookmarks. If a browser could learn my habits, as Safari, is doing and creating bookmarks, I would enjoy that.
- Maybe a better way to visualize more bookmarks at one glance
- folder or tree stucture

2. Do you have any suggestions on recovering files, documents, and web pages for recovering the context of a task?

   Example Data:

   - I always start with very broad keywords and then begin to narrow it down.
   - Having the ability to restore a file to a previous version, like how Sharepoint does
   - I just need to be better at labeling and organizing.

Figure 5.1 shows an example board with task notes clustered onto it.

5.2.4 Study Procedures

For our proper user study, procedure was similar to the pilot study with a few significant changes. Study sessions were held as groups of three participants each. Prior to a study session, the HoloLens devices were disinfected and placed in three separate rooms. Laptops were configured in each room to have a Zoom call providing a channel of communication between each user. Once participants arrived, they were asked to sign consent forms and then collectively instructed on how to wear the HoloLens, what buttons on the device were for, and the sorts of hand gestures they would need to learn for the tasks. Once any questions

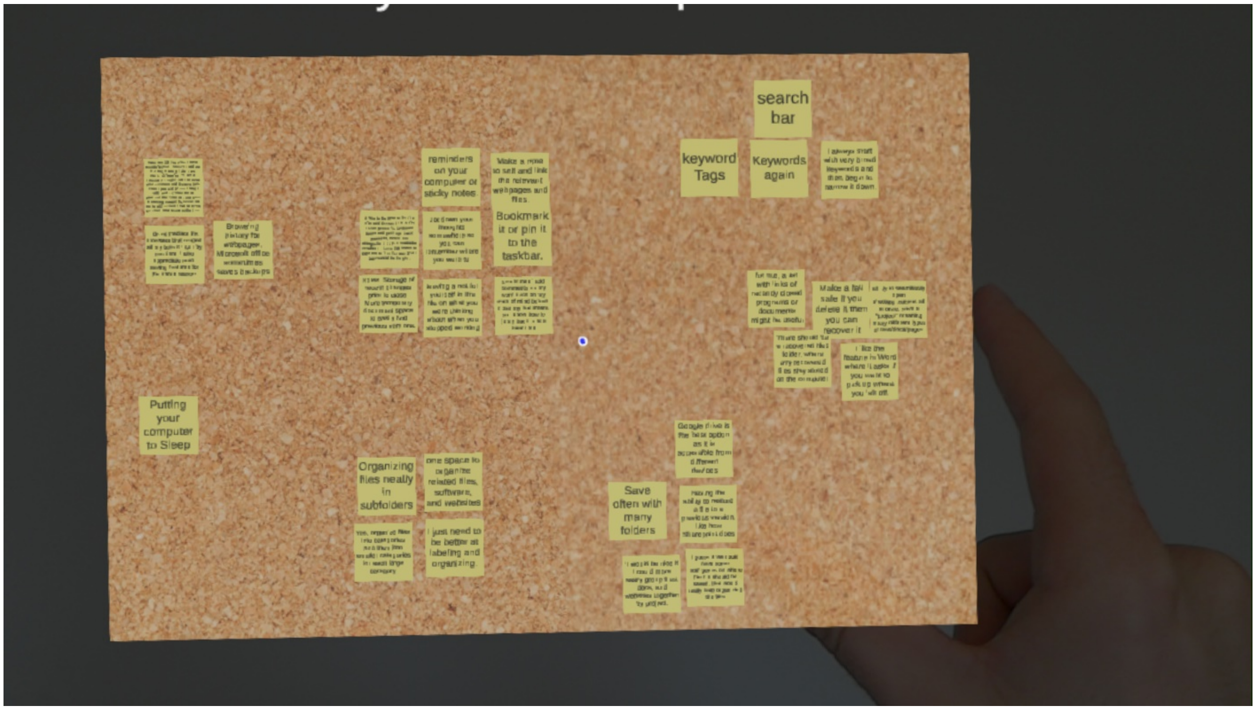


Figure 5.1: Snapshot of a board with task data clustered onto it.

were answered, the participants were separated into the three separate rooms and the study began. Each half of a study session could be broken into three separate phases: training, task, and data collection.

Training

The training phase was intended to familiarize participants with the HoloLens interface as well as the condition of the virtual environment, which could be our baseline cue condition or no cues at all. First, participants were guided through how to open our application on their devices if it was not already open. Then, the appropriate training scenario was loaded, which consisted of the cue condition that would be used in the lengthier task phase as well as a simplistic dataset that was designed to be easily clustered. The participants were instructed to familiarize themselves with the environment and ultimately to cluster the virtual sticky notes onto the board, as they would be doing in the task phase. They were then given 5 minutes to accomplish the simple task, at which point the investigator asked if they had any problems and moved on to the task phase.

Task

After completion of the training phase for a given condition, the investigator changed the configuration of the virtual environment to introduce the new task dataset. The cue condition remained the same as in the training phase. The participants were now given 30 virtual sticky notes to work with rather than 24 from the training session, and the text on the sticky notes was intentionally difficult or ambiguous in order to make them harder to cluster than in the training session. As a result, the participants were given 20 minutes to cluster these notes, with the possibility of finishing early, in which case they were instructed to inform the investigator.

Data Collection

Following the task session, the participants were asked to complete three tasks involved with data collection for the session. First, they were given a sheet of paper with instructions and a screenshot of the in-app virtual board. They were instructed to draw a sketch of the clusters of notes they collaboratively produced and number them from 1 to 6 (if a session had more than 6 clusters, the least significant cluster was ignored). These numbers were referred to in the second data collection task, which was a computer survey that asked the name and description of cluster 1, then cluster 2, and so on. This survey gave participants the opportunity to describe their understanding of the clusters so we could later gauge their level of agreement. The last data collection task was another computer survey. This survey included required System Usability Survey (SUS) questions, as well as short answer questions, that asked about their user experience with the recent environment conditions. Copies of all surveys and data collection documents can be found in the appendix.

5.2.5 Interaction Traces

In addition to the data we collected directly from the participants, our application also recorded various interaction traces from our sessions.

Note Traces

IdeaSpace was configured to record three different interactions related to notes. For each of these interactions, we recorded a timestamp and the user notifying of the event. The interactions we recorded were:

- In/out of view
- In/out of focus
- Move start/end

In or out of view events triggered when a note entered or left a user's view. This was determined using the screen coordinates of the note – if the coordinates were not between 0 and 1, the note was not currently in view. In or out of focus events indicated when a user had placed the HoloLens cursor over a note. We were already tracking this kind of event in order to ensure notes highlighted white when focused, so implementation was trivial. Lastly, move events were triggered whenever a user started and stopped moving a note. For the view and focus events, the only information recorded was the timestamp, the user notifying of the event, and the ID of the note in question. For the move events, however, we recorded all of this information as well as the note's position at start or finish.

Start and End of Task

We configured IdeaSpace to respond to voice commands. Two of these voice commands were commands to record the start and end of a task to ensure we could come up with accurate task durations. Additionally, the start and end commands caused the application to record the start or end state of the task notes, enabling us to reconstruct a view of the board later.

5.3 External Evaluation of the Task Result

Once the user study sessions had completed, we wanted to have a measure that could rate the performance of the users in each of our sessions. As a result, we held external judge sessions where judges were shown select data from the user study sessions and instructed to complete surveys (see A.7) regarding the data. We recruited a total of two (N=2) judges, one female and one male. No personal data was included in the data shown to the judges.

5.3.1 Procedures

External judges sessions were held one person at a time. Upon arrival, the judge was given a paper copy of our consent form and asked to sign it after reading. The judge was then instructed on how to wear the Microsoft HoloLens. The judge then opened IdeaSpace and the investigator spoke a voice command to load the final state of the first task of the first user study session. This data was used to provide context needed to answer some of the survey questions regarding the categorized clusters produced by users from the study task. With the environment state loaded, the judge was then handed copies of the category declaration sheets from the session and was also given access to the survey data from the session that the users used to describe and name the categories. Finally they were asked to fill out a new external judges survey and submit it, at which point the investigator would load state from the next study task end state and offer the user-produced data relevant to it, moving through each in chronological order until the judge had submitted a survey for every task.

5.3.2 Data Collection

As described in the previous section, each judge was instructed to complete an entry of the external judges survey for each completed task (see survey at A.7). The survey questions were 7-point Likert scale and asked the judges to indicate how much they agreed with a statement, from Strongly Disagree to Strongly Agree. These choices were decided upon so that they could be converted to numeric data in a similar way to the SUS and then summed up into a final score.

Chapter 6

Results

In this chapter we discuss the data we collected and the results of our user study.

6.1 Observations

Before offloading of interaction traces from the devices and analyzing the collected data, we were able to make a few observations. We noticed from speaking to participants during and following the study sessions that some had difficulties with manipulating the virtual sticky notes. It seemed likely this was due to unfamiliarity with the HoloLens gesture input. We noticed that in the second task of each of the sessions, the users seemed to have less difficulty, indicating that they had learned how to perform the gestures.

6.2 Task Completion Time

We used recordings of the starts and ends of tasks to calculate task durations. Because we held four experiments with two tasks each, we calculated eight different task durations. We were interested in finding if the cue condition had an effect on the task duration, so we averaged the four durations for each cue condition. The average task duration for the cue condition was 934.13 seconds (standard deviation of 426.26 seconds), while the average duration for the cueless condition was 895.67 (standard deviation of 392.04 seconds). We also checked if there was more of a difference between task sessions that were first in each experiment versus task sessions that were second in each experiment, regardless of cue condition. We found that the average duration of the first tasks in the experiments was 1131.74 seconds (standard deviation of 340.32 seconds), while the average duration of the second tasks in experiments was 698.05 seconds (standard deviation of 308.49 seconds). It took nearly twice as long on average for users to complete their first session of any experiment than to complete the second session. We suspect this gap could be attributed to users acclimating to the environment and the task set before them, learning to complete their second tasks much quicker. Figures 6.1 and 6.2 show comparisons of task completion times by condition and order.

6.2.1 Figure 6.1

In Figure 6.1 the task completion times were graphed by cue condition (with or without cues) as well as by order (whether the task was the first or second of the session) as boxplots. Orange lines are medians and green triangles are means. Note the similarity of the averages and variance of With Cues and Without Cues conditions. Conversely, notice the huge difference between the averages of the First Session tasks versus the Second Session tasks. Additionally, note the greater variance among the First Session tasks than the Second Session tasks. The differences between the First Session tasks and Second Session tasks support our theory that learning effects played a large role.

6.2.2 Figure 6.2

Figure 6.2 breaks the task completion times down by combining cue condition and order to find evidence of interaction between the two. Once again, orange lines are medians and

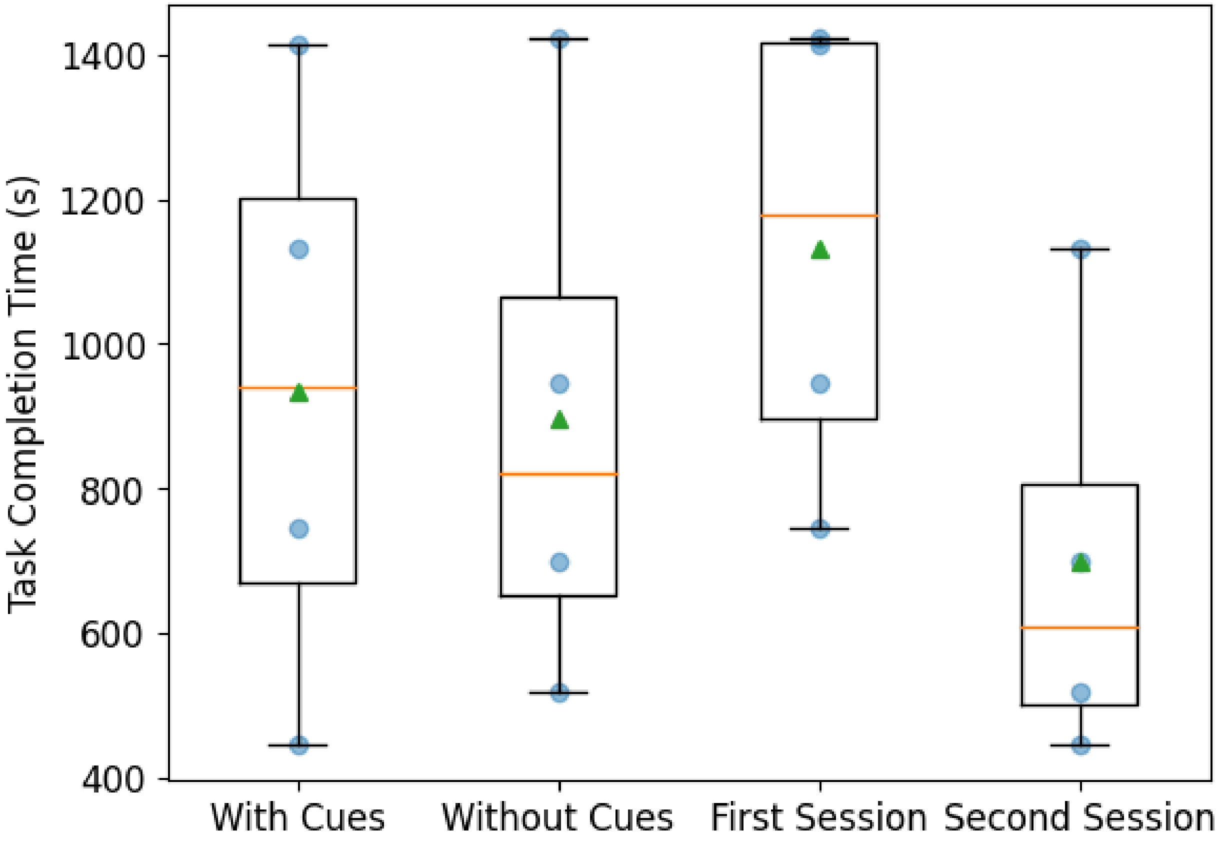


Figure 6.1: Boxplots of task completion times by condition (with or without cues) and order (first or second of session)

green triangles are means. One observation from the figure is that the difference in means between First With Cues and Second With Cues is less than between First Without Cues and Second Without Cues. Additionally, the variance of Second Without Cues is the lowest of the three plots.

6.3 Note Movement Over Time Per User

We chose to record the number of user interactions with virtual sticky notes. Internally, our application recorded every time a user manipulated a note and wrote this information to a

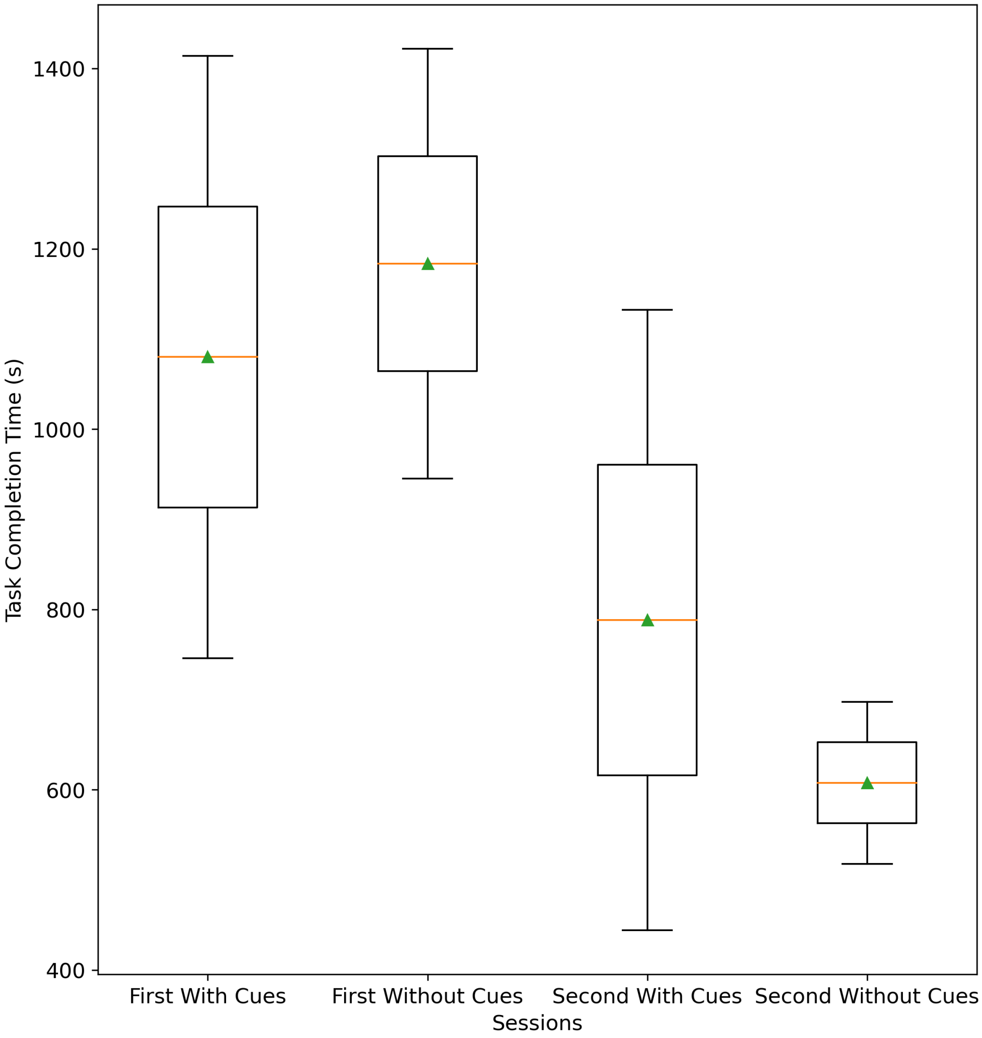


Figure 6.2: Boxplots of task completion times grouped across condition and order

log with a timestamp, the user who initiated the manipulation, the ID of the sticky note, and the starting and ending positions of the note. With the timestamps of these events and records of which users prompted them, we were able to construct step line graphs of the number of manipulation events for each user (see Figures 6.3, 6.4, 6.5, and 6.6).

After creating the step graphs for each session, we compared those in the cue condition with those in the cueless condition to potentially identify patterns. However, we were not able to find a pattern that might support a difference between the cue and cueless conditions for

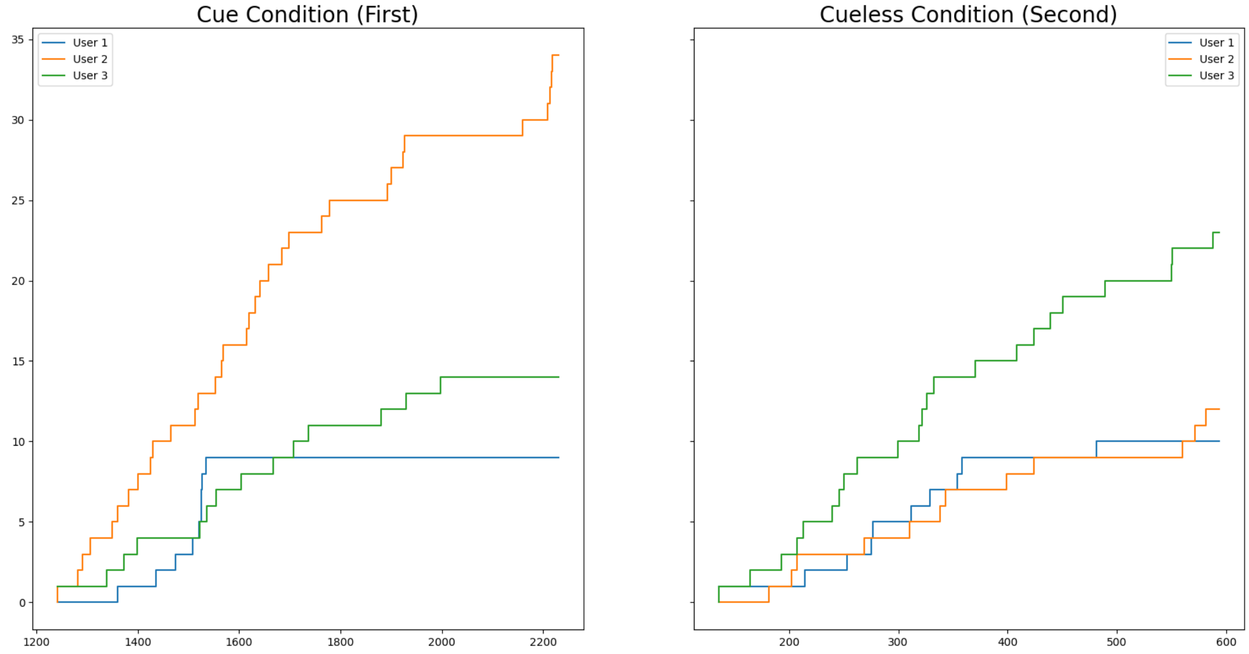


Figure 6.3: Step graph for sums of note controls over time by user and condition for session 1.

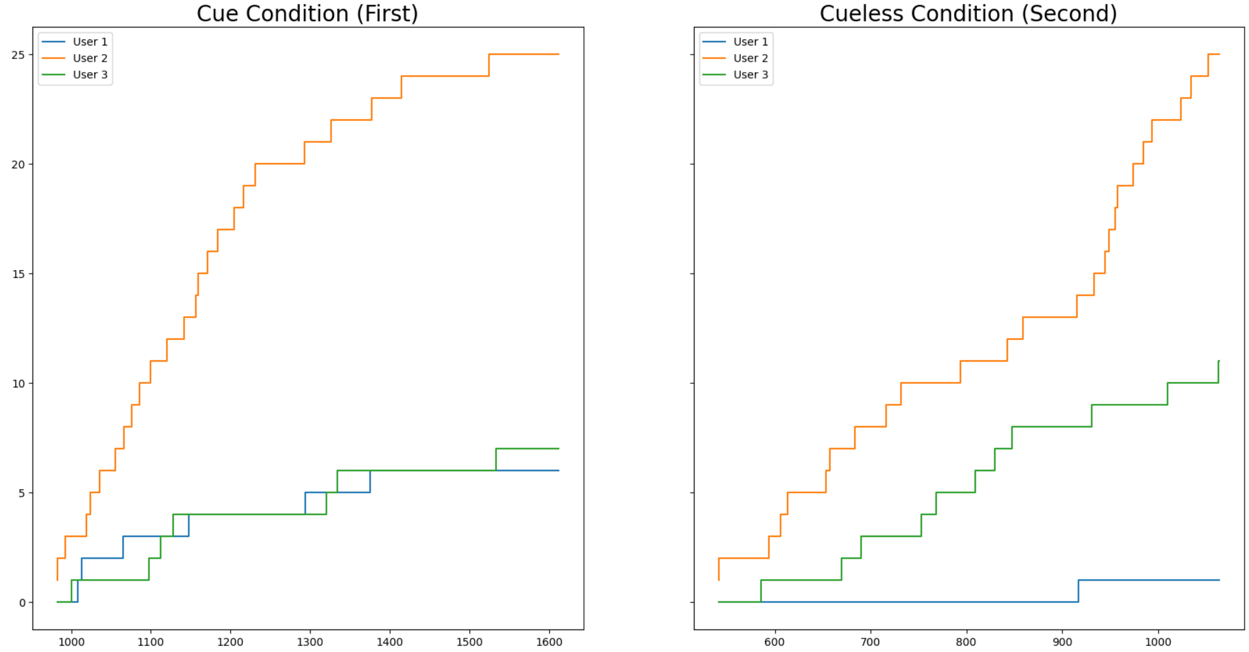


Figure 6.4: Step graph for sums of note controls over time by user and condition for session 2. The flat line in the cueless condition for User 1 was a user giving up trying to move notes.

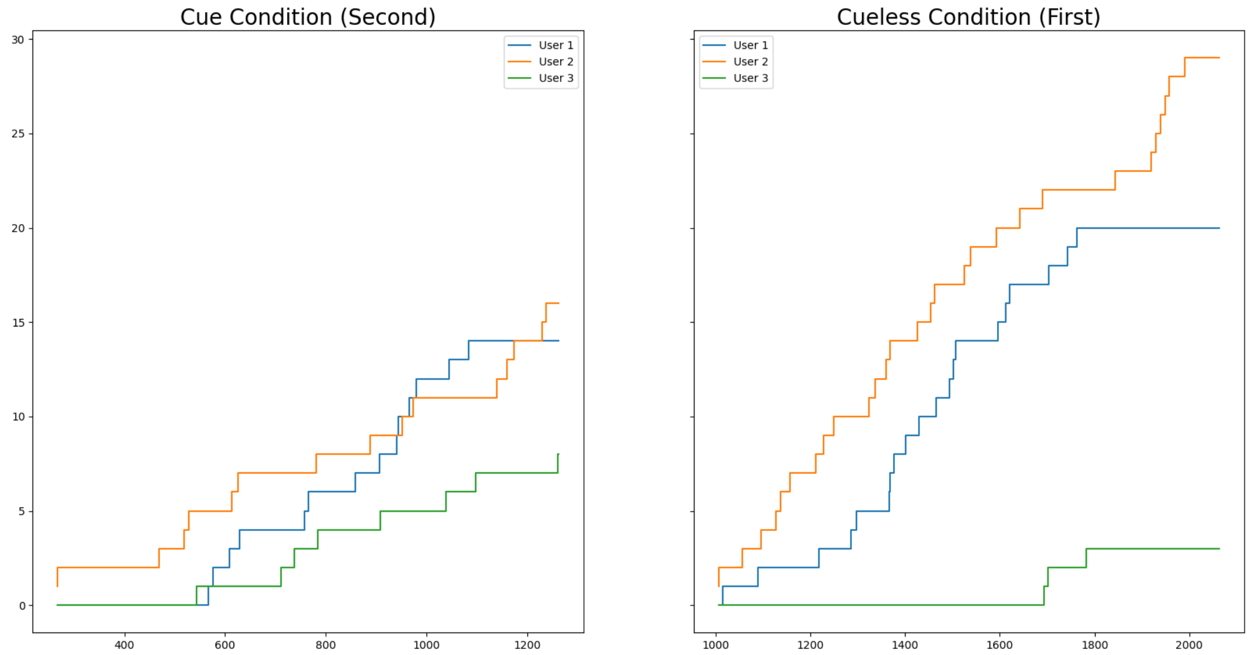


Figure 6.5: Step graph for sums of note controls over time by user and condition for session 3.

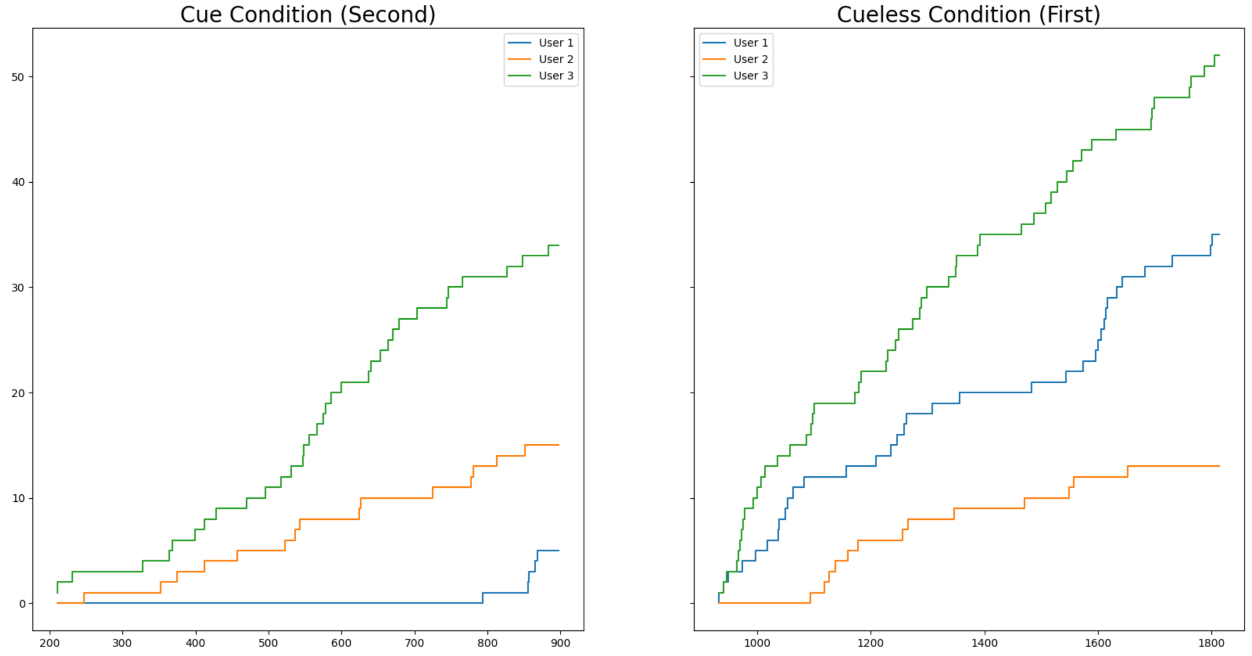


Figure 6.6: Step graph for sums of note controls over time by user and condition for session 4. The flat line in the cue condition for User 1 was a user not being present in the environment until their device battery had charged.

this metric. We did note, however, that in some of the graphs a user's participation was very low. For example, in one graph a user only manipulated a single note, and in a few others the count for some users did not rise above five, with other users in these sessions moving thirty or more notes. Furthermore, we noted that if a task was first in a session, its step graph had less even note movements between users than the second in a session, indicating users learning to move notes who had difficulty in the first task. Users in the second tasks also seemed to used less total note movements, further supporting our learning effects theory.

6.4 Note Movement Through Space

Additionally, we created plots of trajectories of notes in each session, represented by arrows each leading from initiations of note movements to their completions (see Figures 6.7, 6.8, 6.9, and 6.10). These plots were also color coded by user, with the movements whose users could not be determined instead being represented in black. We noticed that in the first tasks in our sessions, the paths of notes seemed to be segmented, while in the second tasks they were more direct, indicating users had learned how to move them more easily in the second tasks.

In this visual plot, we expected to find evidence of users working in specific territories, such as what was shown in Scott et al. [21]. However, it was difficult to tell if this was the case. Because users could work in the same virtual space, perhaps these territorial effects were less prevalent than they could have been.

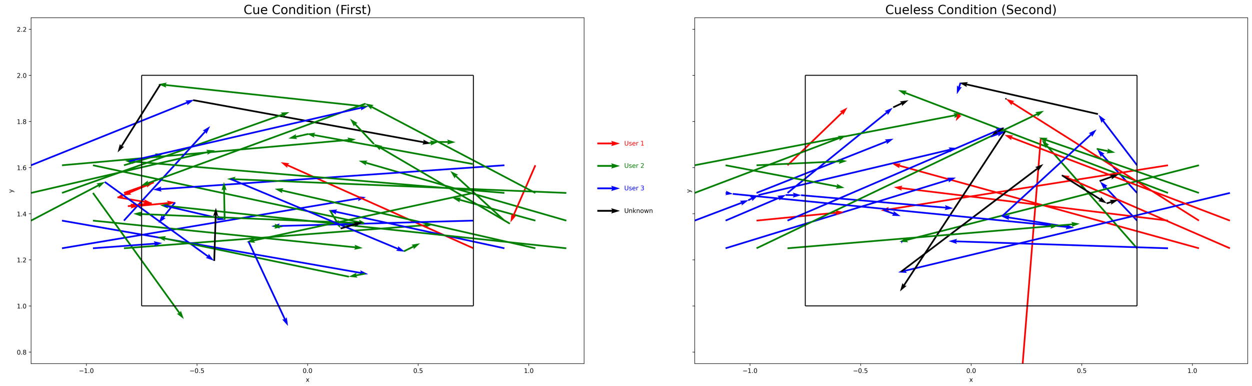


Figure 6.7: Arrow plot of the trajectories of notes over the course of each task in session 1, color coded by user. The Unknown code refers to ambiguous movements, where multiple users could have initiated

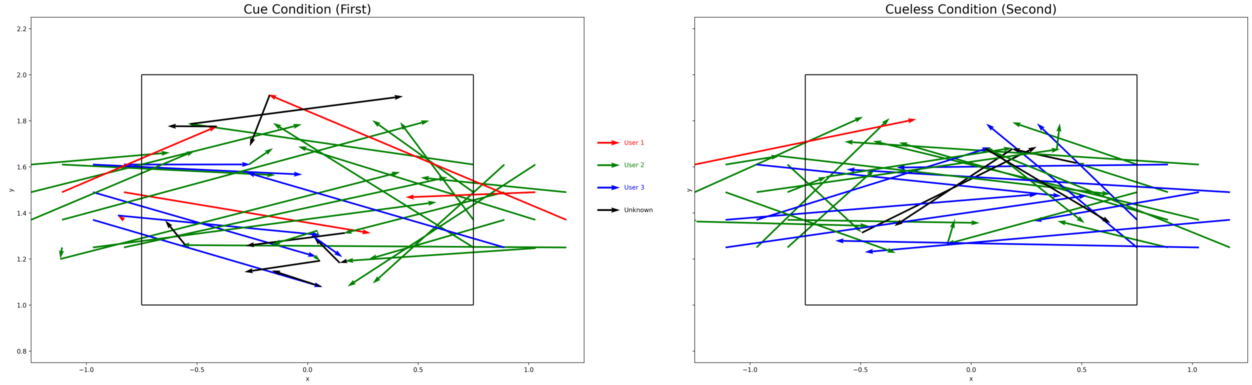


Figure 6.8: Arrow plot of the trajectories of notes over the course of each task in session 2, color coded by user. The Unknown code refers to ambiguous movements, where multiple users could have initiated

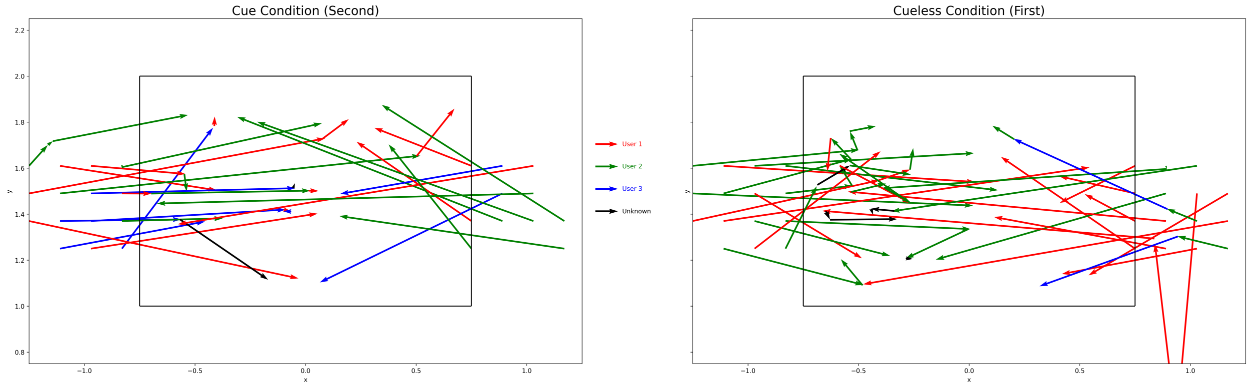


Figure 6.9: Arrow plot of the trajectories of notes over the course of each task in session 3, color coded by user. The Unknown code refers to ambiguous movements, where multiple users could have initiated

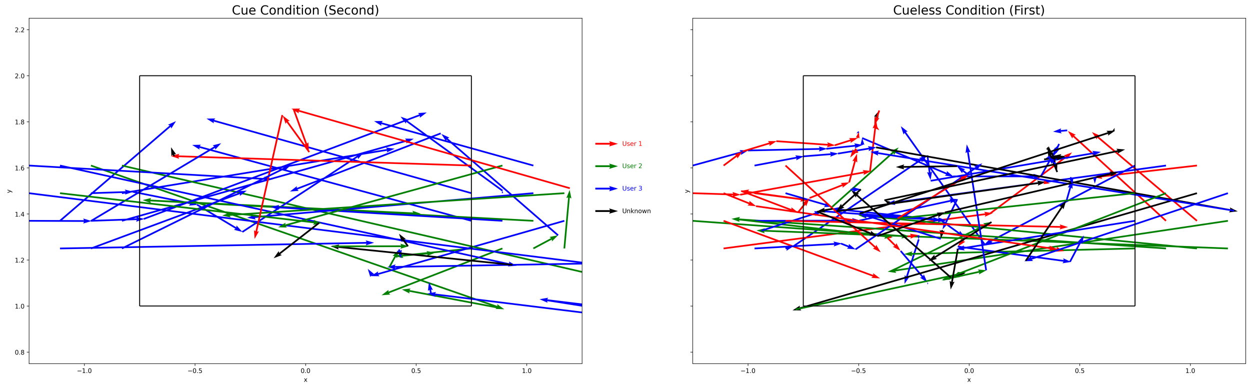


Figure 6.10: Arrow plot of the trajectories of notes over the course of each task in session 4, color coded by user. The Unknown code refers to ambiguous movements, where multiple users could have initiated

6.5 Collaboration Efficiency

Because each note movement event recorded not just the user initiating the event, but also the note being moved, we were able to collect a list of users that interacted with each note in a session. Using this list, we could then count up the number of notes that had interactions from one, two, or three unique users. We felt this could be an interesting metric as an indication of efficiency of collaboration, where a lower number of unique user interactions for notes could indicate more efficient collaboration and a higher number could indicate less. After graphing this metric (see Figures 6.11, 6.12, 6.13, and 6.14), however, we were not able to identify a consistent pattern relating to the cue condition. In some cue condition sessions, there were more notes that had three unique users interact with them over the course of the experiment, and in some there were no notes with three unique users. We did, however, notice a pattern regarding the order of the conditions. Whichever condition came first seemed to have more notes with two or three unique users than the next. This is likely an indication of users learning to move notes more efficiently from one task to the next.

6.6 External Quality Judgements

To obtain a metric of quality, we started first with the responses to the surveys from the external judges. We parsed the choices into numeric values in a similar way to SUS. We took the sum of these values for each session task. For the cue conditions, this process yielded values of 29.5, 32.5, and 32.5. For the cueless conditions, it yielded 30.5, 30.5, and 32.5. Due to a partial loss of data, the quality data consisted of two first tasks and one second task for the cue data and one first tasks and two second tasks for the cueless data. Second tasks tended to be completed faster, regardless of cue condition, so we did not average these

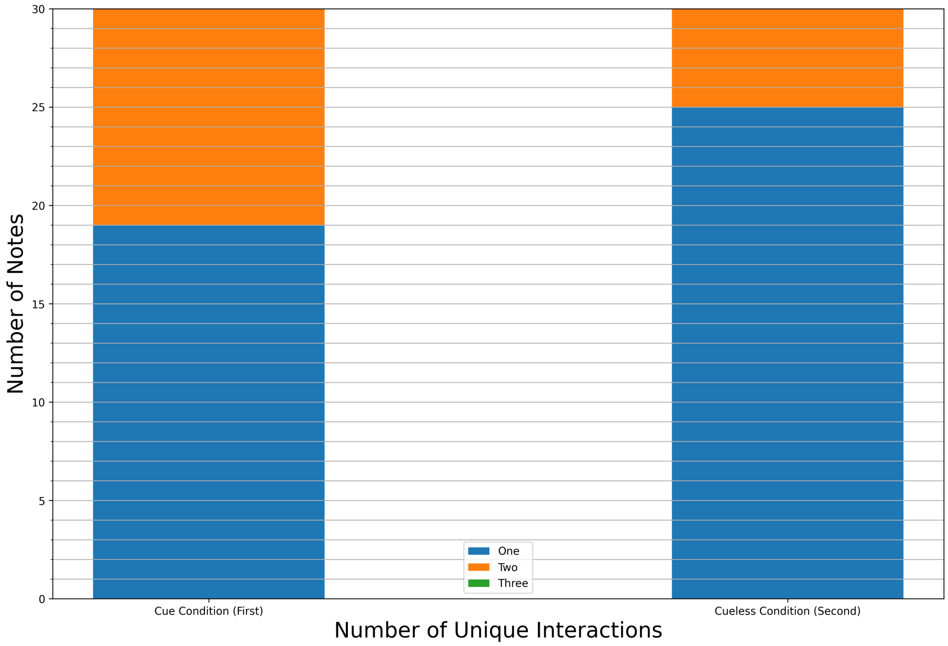


Figure 6.11: Stacked bar chart of number of notes moved by one, two, or three users, compared by condition for session 1.

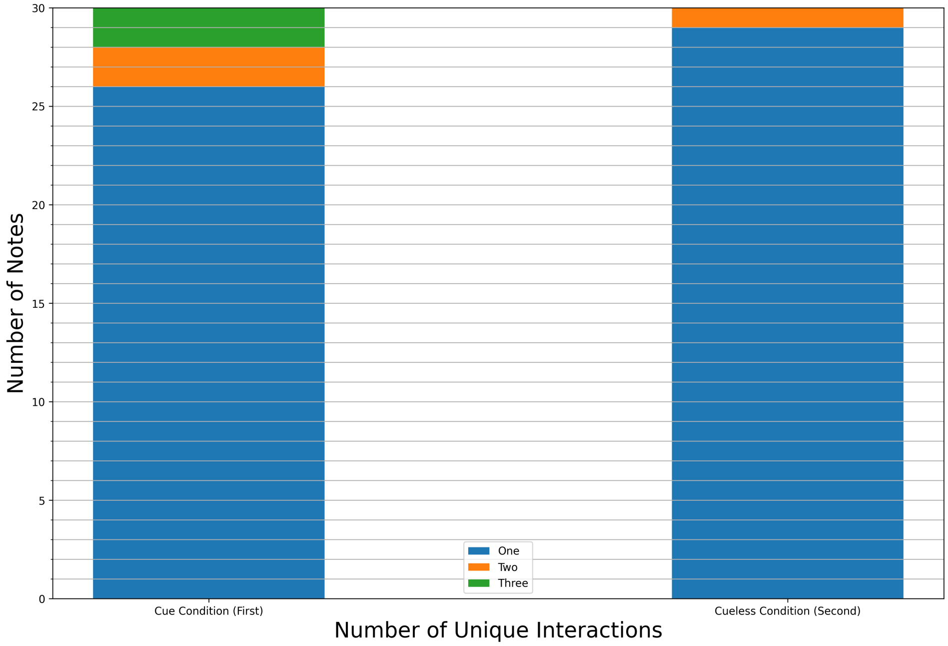


Figure 6.12: Stacked bar chart of number of notes moved by one, two, or three users, compared by condition for session 2.

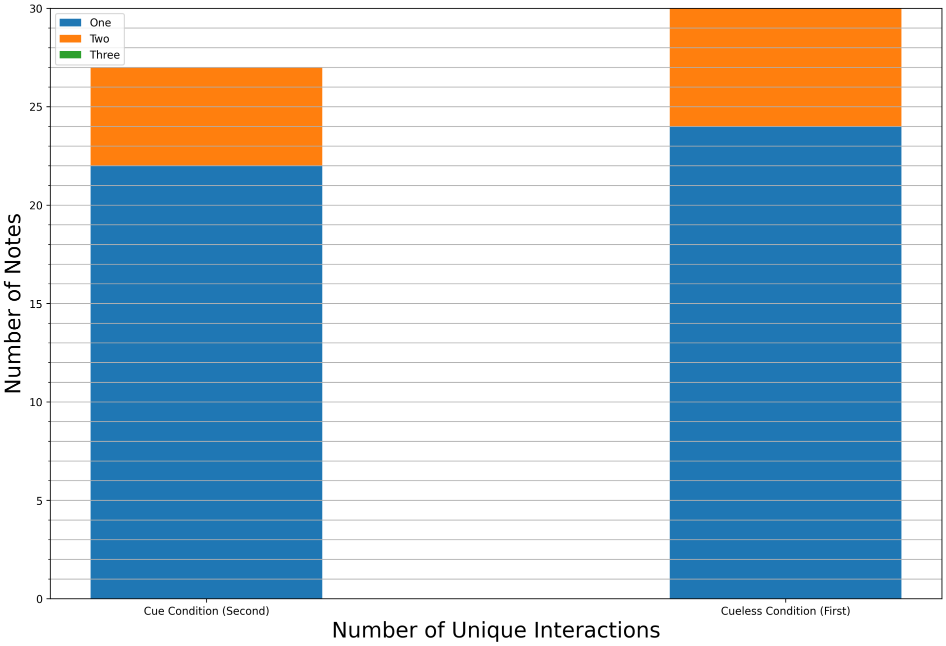


Figure 6.13: Stacked bar chart of number of notes moved by one, two, or three users, compared by condition for session 3. In the cue condition, the users never interacted with three notes, hence the gap at the top.

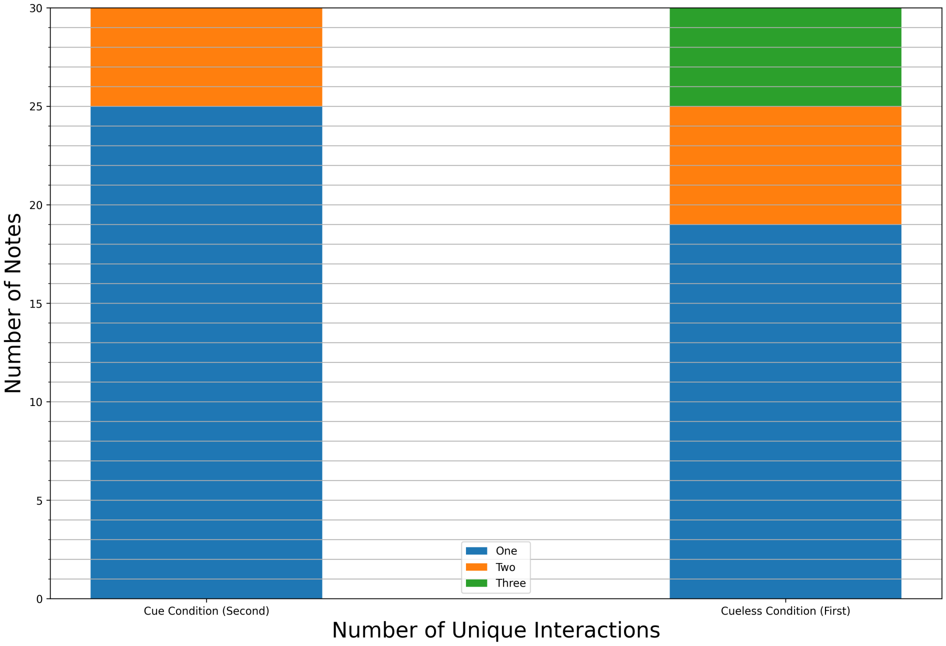


Figure 6.14: Stacked bar chart of number of notes moved by one, two, or three users, compared by condition for session 4.

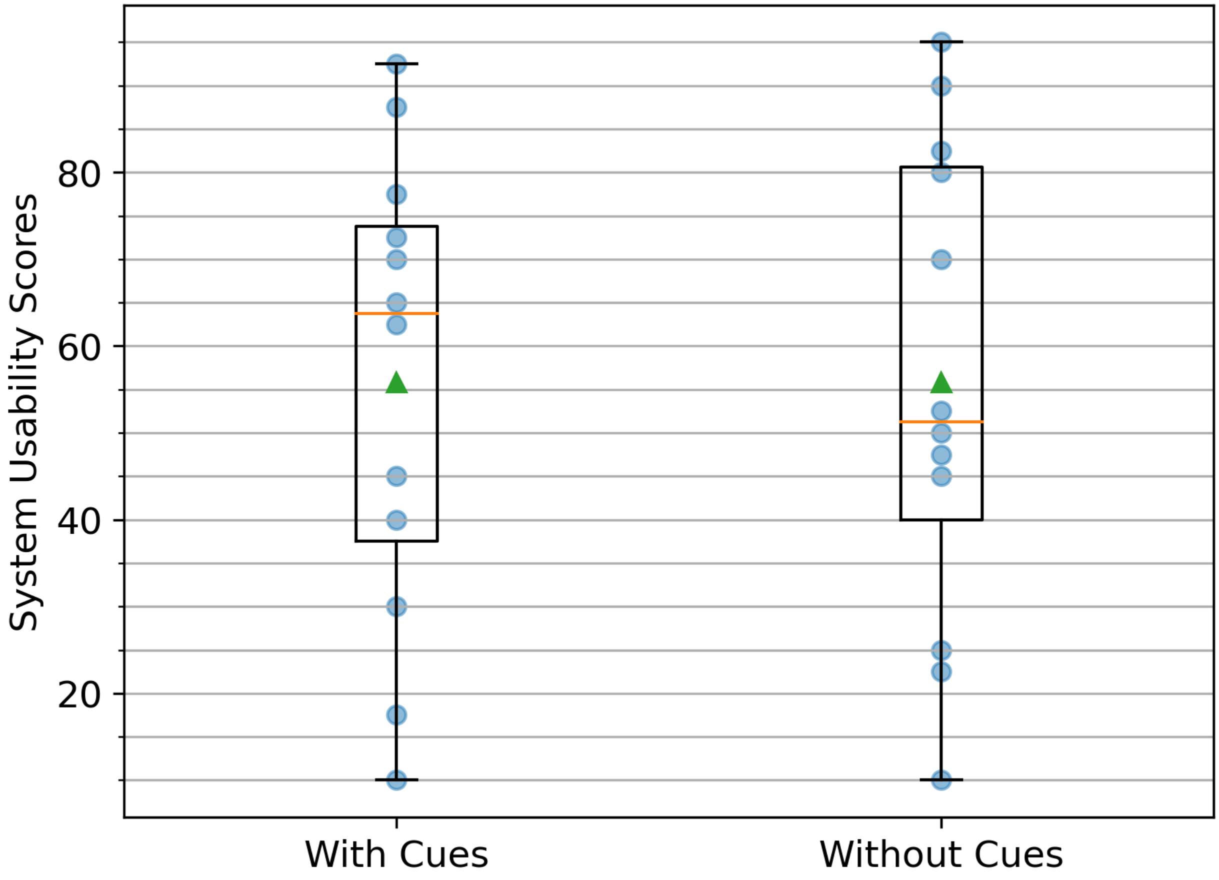


Figure 6.15: Boxplots of SUS scores by condition. Orange lines indicate medians, green triangles indicate means

values as we felt they would be misleading due to the order/learning effect bias.

6.7 System Usability Survey Data

As a part of the survey data we collected from participants during the experiment, we asked ten multiple choice questions known as the System Usability Survey (SUS). The SUS asked users to rate our system by how much they agreed with ten statements about the system, and it yielded a score on a scale from 0-100. After collecting survey responses from all twelve participants, we averaged the ratings of the cue condition and cueless condition. Both cue conditions scored 55.83. In SUS scoring, 68 is considered an average score, so both conditions scored below average. Something we noted was the wide range of scores, which we visualized as a boxplot in Figure 6.15. The range seemed to indicate some users thought IdeaSpace was easy to use, and others thought it was difficult, so we investigated further.

6.7.1 Low Scores

Participants 4, 6, and 10 gave SUS scores of 17.5/22.5, 10/10, and 30/25, respectively, for the cue/cueless conditions. Their scores comprised the lowest scores from our participants, so we were interested in their answers to our open-ended questions. In both tasks, participant 4 listed note movement as their biggest challenge (“I couldn’t grab the sticky notes”, “The grabbing was not working”). Similarly, participant 6 responded “how frustrating it was that the same gesture only worked sometimes, and I didn’t know what made it work”, and participant 10 said “The interaction; it didn’t seem to work most of the time.” Participant 10 also mentioned “Viewing the AR iamge or interface was difficult given the weight of the headset,” perhaps suggesting they had additional difficulty with the headset itself. Participant 6 advised using some external device for input rather than hand gestures (“Having some kind of handheld device to select and point would have made it much less frustrating”, “...having some handheld device to select and move objects would make it less frustrating.”) when asked about tools that might have made the tasks easier, making it clear that note movement via hand gestures was a continuing problem for them. Participant 10 suggested using Miro’s virtual whiteboard instead (“Using Miro, a virtual whiteboard”), and participant 4 did not provide answers to the question, perhaps indicating they were not sure how they could have improved their performance in the task. All in all, these three participants seemed to have had a very difficult time learning our remote affinity diagramming system – specifically how to move notes consistently. This seems to account for their particularly low

SUS scores.

6.7.2 High Scores

In contrast to 4, 6, and 10, participants 12, 7, and 5 provided the highest SUS scores of 87.5/95, 92.5/90, and 77.5/80, respectively, for the cue/cueless conditions. Following their first task (cueless), participant 12 responded that their greatest challenge was “Getting the first few items to move”, but after the second task (cue) they only mentioned “Locating the notes being mentioned”, perhaps indicating that they learned how to better move notes across tasks. Answering the same question, participant 7 did not even mention note movement in their first task (cue): “Collaborating with others. It was hard to talk about certain notes with just words. Wish there was an in-app feature to know what my collaborator is looking at.” Participant 5 said “The most challenging part of using this tool was grabbing the notes and getting them to the right place on the board,” indicating they had some trouble in their second (cueless) task, however they did not mention note grabbing in answers following their first task, except to say that “The small field-of-view made it difficult to control the notes effectively.” Surprisingly, when asked for suggestions that might have made their task easier, participant 7 wrote “Visual cues that would show the collaborator’s position and their head direction” following their first (cueless) task. The same participant suggested removing the hand visuals for less clutter after their second (cue) task. Participant 12 responded “Controller but not necessarily” for their first task with the cueless condition, and later did not list any ideas for their second (cue) task. Participant 5 appreciated the collaborative cues, saying “Visual cues are definitely the way to go” when similarly asked for ideas for helpful tools. A common trend through the answers of participants 12, 7, and 5 seems to be an easier time with moving notes from the start, which improved even further in their second task, resulting in a far more positive experience than participants 4, 6, and 10. This all further emphasizes a need for a longer and more thorough training period in our study, which we elaborate upon in the Discussion chapter.

6.8 Qualitative Survey Data

After users completed the SUS portion of our post-task survey, the next portion asked a number of qualitative questions, giving participants the opportunity to leave a short answer response. Although the short-answer questions were optional, most participants took the time to leave responses for them. When we looked through their responses, we noticed some common threads and complaints. Visit Section A.5 in the appendix to see all the survey questions.

6.8.1 Moving Notes Was Difficult

We noticed most users reported moving notes as being difficult ("I couldn't grab the sticky notes," "The most challenging part of the task was picking up the notes and moving them around the board," "moving the notes"), to the point where some users even recommended using a controller or pointer device rather than our implementation ("Having some kind of handheld device to select and point would have made it much less frustrating," "Controller but not necessarily"). Because HoloLens devices use cameras on the front to detect hand gestures, it is possible the device was having trouble determining when users were performing such gestures due to the lab environment or users performing gestures incorrectly. The latter seems most likely given the differences between first and second conditions in each session.

6.8.2 Cues Helped the Task

To the question “Do you feel the visible cues (including gaze rays, highlights) helped with the task?”, most users answered yes. Some users also noted that the cues helped them when they tried to point things out to others. In response to the question “What was the most challenging in collaboration and communicating with other people, especially regarding referencing virtual objects?” following the cueless condition, one user wrote “We had to come up with ways to make sure we are taking about the same thing. Wish there was an in-app to reference the objects.”

6.8.3 Potential Evidence of Clutter

Although users seemed to unanimously appreciate the cue condition, a few also noted that they were distracting at times. For example, in response to the question “Was having cues at all distracting to the collaboration or focusing on your own tasks?” a user wrote “Yes. Sometimes I was trying to read a sticky note and someone else’s lines were covering up the words.” Another user wrote simply “distracting”. Others disagreed, but these responses seemed to provide evidence that some of the cues present in our cue condition could get in the way of users at times.

Chapter 7

Discussion

In this chapter, we discuss the results from the user study and reflect on the limitations of this thesis.

7.1 Results and Research Questions

At the beginning of this project, we set out to answer three research questions:

RQ1: To what extent can AR-based affinity diagramming be usable?

RQ2: How do awareness cues or affordances for remote collaborative AR effect performance and perceived quality of collaboration in a collaborative affinity diagramming task with multiple users?

For our first question, our study determined that our system was not as usable as it should have been. The SUS scores of both conditions scored below the average of 68. The qualitative feedback from participants also indicated that our system could be improved. Note movement, perhaps the most important part of the task, was the aspect that users struggled the most with according to their feedback. This difficulty likely played a significant role in the other results, as well as the learning effects as they overcame this obstacle.

Regarding our second research question, our quantitative metrics did not provide persuasive evidence either way. When it came to task completion time, note movement, and our measures of efficiency of collaboration, there were no significant differences between the cue and cueless conditions. Similarly, the SUS scores were too similar to conclude one was better than the other. However, our qualitative survey data seemed to indicate that users preferred having the cues to not having them, especially the cues indicating where other users were looking. These cues enabled users to indicate a particular sticky note without having to describe it verbally. Additionally, we noticed a difference on every metric we recorded based on which condition was first or second in a session. This seemed to indicate strong learning effects, which we address further in the Limitations section.

7.2 Implications

Most of the users in our study had trouble getting the device to detect their hand gestures, and a few even expressed a desire for a controller rather than having to rely on the gestures. From this, we were able to conclude that the training portions of our user study were not sufficient for users to have learned how to use our drag-and-drop movement technique. An implication from our study, then, was that it is important in a study for participants to demonstrate proficiency with a tool before using it for a task. We further discuss this and changes we should have made in Section 7.3.1.

7.3 Limitations

We acknowledge that our project and study had a number of limitations, which are likely the reasons for such inconclusive results and the gaps between first and second tasks. We list each of these in the following section and try to address how they might have interfered with the application and our user study.

7.3.1 Learning Effects

Throughout the task during our user study, we noticed users were still having trouble moving notes, even after the training. This observation was corroborated by our collected data, which showed some users barely interacted with notes in tasks, and others resorted to many small movements of notes in order to get them in the desired locations. Our data also showed that users were able to move notes more easily in their second task, regardless of condition. This indicates a learning effect across tasks and shows that users had not mastered the manipulation methods before going into their first tasks. This limited our analysis and could have been mitigated by allowing for longer training times, testing the users until they demonstrated proficiency before moving on to the task, or making the training task individual so that all users were given a sufficient chance to try out the input methods.

It is worth discussing what exactly users struggled to learn as they used our system. From our observations, it seemed specifically that users could not get their gestures to be recognized in the first place, or at least not consistently. We believe this has to do with the fact that three HoloLens gestures (Tap, Hold, and Drag) all start out the same. Similar to a double click gesture on a computer, it is likely that there is a delay where the HoloLens waits for another event to occur following the start of a gesture in order for it to determine which gesture a user is trying to use. Thus, for a Drag gesture the HoloLens needs to recognize the start of a gesture and then wait for the detected hand to move a certain amount in a certain time frame to differentiate a Drag gesture from a Hold gesture. We suspect many users when initially attempting the Drag gesture were either shaping their hands incorrectly or trying to move their hands too quickly or too slowly, and in later tasks they learned the extent of the gesture detection space as well as the timing needed in order for their gestures to be correctly detected.

7.3.2 Co-presence

Another potential limitation of our project and study was a lack of focus on co-presence. Co-presence is the feeling of being present in the same space [19]. Piumsomboon et al. [19] noted that studies have shown gaze cues can help increase co-presence of users. However, we did not attempt to measure or ask users about their sense of co-presence in our user study.

Because we did not directly measure co-presence, we can only guess about it based on observations during the study. One potential indicator of co-presence would have been users speaking in the direction of other users' head visuals during the study. However, we did not observe this and suspect the Zoom calls interfered with it. Because we had a Zoom call set up on laptops for each user, users spoke toward the laptops during task, from where they were hearing other user audio. It is likely this interference could have been fixed with spatial audio so that users would hear other user audio from their locations in the virtual world.

7.3.3 Logging Errors

We noticed following the user study sessions that our logging code had not worked as expected. The first problem was that the note movement events had all been logged as coming from whichever device had started the application first in the session. Fortunately, we had also recorded user focus events, when users placed their cursors over notes, and a focus event necessarily preceded a move event. As a result, we were able to recover the vast majority of the move event senders, excluding those that were indiscernible from our data.

7.3.4 Small Sample Size

A likely limitation of our study is the small sample size. We were only able to recruit 12 participants to participate in the four experiments. As a result, it is entirely possible that our study population is not representative of the whole.

Another product of the small sample size was likely the over representation of learning effects. Only four of the participants reported that they were moderately or more familiar with AR headsets and the Microsoft HoloLens. We noticed from comparisons of task completion time and other metric data that participants tended to perform better and more quickly in their second task than their first, indicating that they were likely getting used to the HoloLens and the task and got better for the second task. If we had been able to recruit more participants that were already familiar with AR technology, we might have returned more interesting or meaningful results from our experiment.

7.3.5 Covid-19

Our recruiting and experiments took place during the Covid-19 pandemic. While this did not significantly affect the study procedure beyond sanitizing devices and socially distancing participants, it is likely that this diminished the number of people willing to participate.

Chapter 8

Conclusion

Our project sought first to create an AR application that supported common brainstorming tasks. The goal was to try to leverage the immersive advantages of AR and the technical advantages of digital brainstorming in order to develop the basis for a software that could improve brainstorming, particularly by enabling these advantages to be utilized remotely. We found that our system was not as usable as it could be and that it leaves significant room for improvement.

Our project also sought to investigate the effects AR collaborative cues have on brainstorming tasks. For this, we carried out a user study where participants completed clustering tasks with and without modern collaborative cues. We did not observe a severe difference between the cue and cueless conditions, which could be due to a variety of factors, but mostly the learning effects of users from task to task.

Bibliography

[1] Features overview | photon engine. URL https://doc.photonengine.com/en-us/pun/current/getting-started/feature-overview.

[2] Doug Bowman and Larry Hodges. An evaluation of techniques for grabbing and manipulating remote objects in immersive virtual environments. *Symposium on Interactive 3D Graphics*, 182, 09 1999. doi: 10.1145/253284.253301.

[3] Susan E. Brennan, Xin Chen, Christopher A. Dickinson, Mark B. Neider, and Gregory J. Zelinsky. Coordinating cognition: The costs and benefits of shared gaze during collaborative search. *Cognition*, 106(3):1465–1477, 2008. ISSN 0010-0277. doi: https://doi.org/10.1016/j.cognition.2007.05.012. URL https://www.sciencedirect.com/science/article/pii/S0010027707001448.

[4] Sarah D'Angelo and Andrew Begel. *Improving Communication Between Pair Programmers Using Shared Gaze Awareness*, page 6245–6290. Association for Computing Machinery, New York, NY, USA, 2017. ISBN 9781450346559. URL https://doi.org/10.1145/3025453.3025573.

[5] Sarah D'Angelo and Darren Gergle. *Gazed and Confused: Understanding and Designing Shared Gaze for Remote Collaboration*, page 2492–2496. Association for Computing Machinery, New York, NY, USA, 2016. ISBN 9781450333627. URL https://doi.org/10.1145/2858036.2858499.

[6] Barrett Ens, Joel Lanir, Anthony Tang, Scott Bateman, Gun Lee, Thammathip Piumsomboon, and Mark Billinghurst. Revisiting collaboration through mixed reality: The evolution of groupware. *International Journal of Human-Computer Studies*,

131:81–98, 2019. ISSN 1071-5819. doi: https://doi.org/10.1016/j.ijhcs.2019.05.011. URL `https://www.sciencedirect.com/science/article/pii/S1071581919300606`. 50 years of the International Journal of Human-Computer Studies. Reflections on the past, present and future of human-centred technologies.

[7] S. Feiner and O. Oda. 3d referencing techniques for physical objects in shared augmented reality. In *2012 IEEE International Symposium on Mixed and Augmented Reality (ISMAR 2012)*, pages 207–215, Los Alamitos, CA, USA, nov 2012. IEEE Computer Society. doi: 10.1109/ISMAR.2012.6402558. URL `https://doi.ieeecomputersociety.org/10.1109/ISMAR.2012.6402558`.

[8] Steffen Gauglitz, Benjamin Nuernberger, Matthew Turk, and Tobias Höllerer. World-stabilized annotations and virtual scene navigation for remote collaboration. In *Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology*, UIST '14, page 449–459, New York, NY, USA, 2014. Association for Computing Machinery. ISBN 9781450330695. doi: 10.1145/2642918.2647372. URL `https://doi.org/10.1145/2642918.2647372`.

[9] Kunal Gupta, Gun A. Lee, and Mark Billinghurst. Do you see what i see? the effect of gaze tracking on task space remote collaboration. *IEEE Transactions on Visualization and Computer Graphics*, 22(11):2413–2422, 2016. doi: 10.1109/TVCG.2016.2593778.

[10] Keita Higuchi, Yinpeng Chen, Philip A. Chou, Zhengyou Zhang, and Zicheng Liu. *ImmerseBoard: Immersive Telepresence Experience Using a Digital Whiteboard*, page 2383–2392. Association for Computing Machinery, New York, NY, USA, 2015. ISBN 9781450331456. URL `https://doi.org/10.1145/2702123.2702160`.

[11] Allison Jing, Kieran May, Gun Lee, and Mark Billinghurst. Eye see what you see: Exploring how bi-directional augmented reality gaze visualisation influences co-located

symmetric collaboration. *Frontiers in Virtual Reality*, 2:79, 2021. ISSN 2673-4192. doi: 10.3389/frvir.2021.697367. URL https://www.frontiersin.org/article/10.3389/frvir.2021.697367.

[12] Dan Kallen. Hololens 2 vs. hololens 1: What's new?, 2020. URL https://hartfordrents.com/blog/hololens-2-vs-hololens-1/.

[13] Jiro Kawakita. The original kj method. *Tokyo: Kawakita Research Institute*, 5, 1991.

[14] Seungwon Kim, Gun Lee, Nobuchika Sakata, and Mark Billinghurst. Improving co-presence with augmented visual communication cues for sharing experience through video conference. In *2014 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pages 83–92, 2014. doi: 10.1109/ISMAR.2014.6948412.

[15] Yuan Li, Donghan Hu, Boyuan Wang, Doug A. Bowman, and Sang Won Lee. The effects of incorrect occlusion cues on the understanding of barehanded referencing in collaborative augmented reality. *Frontiers in Virtual Reality*, 2:76, 2021. ISSN 2673-4192. doi: 10.3389/frvir.2021.681585. URL https://www.frontiersin.org/article/10.3389/frvir.2021.681585.

[16] Jingxian Liao and Hao-Chuan Wang. Gestures as intrinsic creativity support: Understanding the usage and function of hand gestures in computer-mediated group brainstorming. *Proc. ACM Hum.-Comput. Interact.*, 3(GROUP), December 2019. doi: 10.1145/3361124. URL https://doi.org/10.1145/3361124.

[17] Stephan Lukosch, Mark Billinghurst, Leila Alem, and Kiyoshi Kiyokawa. Collaboration in augmented reality. *Computer Supported Cooperative Work (CSCW)*, 24, 12 2015. doi: 10.1007/s10606-015-9239-0.

[18] Liang Men and Nick Bryan-Kinns. Lemo: Exploring virtual space for collaborative creativity. 05 2019. doi: 10.1145/3325480.3325495.

[19] Thammathip Piumsomboon, Arindam Dey, Barrett Ens, Gun Lee, and Mark Billinghurst. The effects of sharing awareness cues in collaborative mixed reality. *Frontiers in Robotics and AI*, 6:5, 2019. ISSN 2296-9144. doi: 10.3389/frobt.2019.00005. URL `https://www.frontiersin.org/article/10.3389/frobt.2019.00005`.

[20] VICKY L. PUTMAN and PAUL B. PAULUS. Brainstorming, brainstorming rules and decision making. *The Journal of Creative Behavior*, 43(1):29–40, 2009. doi: https://doi.org/10.1002/j.2162-6057.2009.tb01304.x. URL `https://onlinelibrary.wiley.com/doi/abs/10.1002/j.2162-6057.2009.tb01304.x`.

[21] Stacey D. Scott, M. Sheelagh T. Carpendale, and Kori Inkpen. Territoriality in collaborative tabletop workspaces. In *Proceedings of the 2004 ACM Conference on Computer Supported Cooperative Work*, CSCW '04, page 294–303, New York, NY, USA, 2004. Association for Computing Machinery. ISBN 1581138105. doi: 10.1145/1031607.1031655. URL `https://doi.org/10.1145/1031607.1031655`.

[22] Raymond Scupin. The kj method: A technique for analyzing data derived from japanese ethnology. *Human Organization*, 56, 06 1997. doi: 10.17730/humo.56.2.x335923511444655.

[23] Anthony Tang, Michel Pahud, Kori Inkpen, Hrvoje Benko, John C. Tang, and Bill Buxton. Three's company: Understanding communication channels in three-way distributed collaboration. In *Proceedings of the 2010 ACM Conference on Computer Supported Cooperative Work*, CSCW '10, page 271–280, New York, NY, USA, 2010. Association for Computing Machinery. ISBN 9781605587950. doi: 10.1145/1718918.1718969. URL `https://doi.org/10.1145/1718918.1718969`.

[24] William Widjaja and Makoto Takahashi. Distributed interface for group affinity-diagram brainstorming. *Concurrent Engineering*, 24(4):344–358, 2016. doi: 10.1177/1063293X16657860. URL `https://doi.org/10.1177/1063293X16657860`.

Appendices

Appendix A

Data Collection Forms and Surveys

A.1 User Study Consent Form

Title of research study: *IRB 21-170 Evaluating Collaborative Cues in AR Affinity Diagramming Tasks*

Principal Investigator: *Sang Won Lee* *sangwonlee@vt.edu* *(540)231-4857*

Other study contact(s):

Doug A. Bowman	*dbowman@vt.edu*	*(540) 231-2058*
Nathaniel Llorens	*lnathaniel@vt.edu*	*(561) 400-4804*

Key Information: The following is a short summary of this study to help you decide whether or not to be a part of this study. More detailed information is listed later on in this form.

We invite you to participate in this Augmented Reality (AR) study to help us testing different collaborative techniques. This work is hoping to make AR more usable and we deeply appreciate your participation.

If you agree to participate, you will perform a set of collaborative tasks in an AR environment with two other study participants. First, you will be shown how to wear and use the Microsoft HoloLens that will be provided. You and the two other participants will complete a short training session (approximately 10 minutes) and then two additional task sessions (approximately 20 minutes each) with a training session and break in between. The training sessions and each of the tasks will involve moving virtual sticky notes into clusters that you and your partners agree upon. After the last task session, you and the other participants will answer some interview questions. The total time commitment for this study is around 100 minutes and you will be compensated with a $18 Amazon gift card for your time.

Detailed Information: The following is more detailed information about this study in addition to the information listed above.

Who can I talk to?

If you have questions, concerns, or complaints, or think the research has hurt you, talk to the research team at *lnathaniel@vt.edu / (561) 400-4804.*

This research has been reviewed and approved by the Virginia Tech Institutional Review Board (IRB). You may communicate with them at 540-231-3732 or irb@vt.edu if:

- You have questions about your rights as a research subject
- Your questions, concerns, or complaints are not being answered by the research team
- You cannot reach the research team
- You want to talk to someone besides the research team to provide feedback about this research

How many people will be studied?

We plan to include about 30 people in this research study.

What happens if I say yes, I want to be in this research?

The study will take place in the Usability Lab 133 (D, E, F) (102 McBryde Hall) and will take approximately 100 minutes. When you arrive, you will be greeted and asked to read and sign the informed consent form after your questions (if any) are answered. Then you will be asked to clarify you have normal vision. Next you will be provided with written or verbal instructions for the experiment, and familiarized with the lab and the equipment they will be using. You will wear an augmented reality (AR) headset such as the Microsoft Hololens. Using the devices, you will then complete a collaborative affinity diagramming task with two other users. Tasks will involve physical movements including looking around the environment, manipulating virtual sticky notes, and organizing them into groups. As explained above, there will be three sessions: 1 short training session (10 minutes) to familiarize you and the other users, a task session, a short break, another training session, and a last task session (20 minutes for each task session). After the tasks are completed, you will be interviewed about the entire experience. The entire session will be audio and video recorded via the Zoom call.

To thank you for your participation, we will give out a $18 electronic gift card at the end of the study.

What happens if I say yes, but I change my mind later?

You can leave the research at any time, for any reason, and it will not be held against you.

If you decide to leave the research, contact the investigator so that the investigator can terminate the current experiment session. Any data collected by far will be disregarded and will not be used for further study. There will be no follow-up procedures afterwards.

Withdrawing from the experiment does not affect receiving $12 gift card compensation.

Is there any way being in this study could be bad for me? (Detailed Risks)

Using AR technology can produce symptoms of sickness or discomfort in some users. These symptoms are usually mild, and may include dizziness, nausea, eye strain, headache, or disorientation. During tasks involving physical movement, there is also some risk that you could collide with obstacles in the physical environment.

You will be given the option to take a break or quit the experiment at any time. To mitigate the risk of sickness and discomfort, we will adjust the display properly, keep task sessions short, provide frequent breaks, and ask you after each set of tasks how you are feeling. To mitigate the risk of physical obstacles and cabling, we will clear the area of obstacles, show you where the boundaries of the space are.

What happens to the information collected for the research?

We will make every effort to limit the use and disclosure of your personal information, including research study, only to people who have a need to review this information. We cannot promise complete confidentiality. Organizations that may inspect and copy your information include the IRB, Human Research Protection Program, and other authorized representatives of Virginia Tech.

All data collected during this study will be done so anonymously and stored in a password protected computer. No names, contact information, or any other identifying information will be attached to your responses to an investigator's questions or to your results from the provided tasks. At no time will the researchers release identifiable results of the study to anyone other than individuals working on the project without your written consent. However, you may be asked to perform tasks in a public space and bystanders could recognize you.

If identifiers are removed from your private information or samples that are collected during this research, that information or those samples could be used for future research studies or distributed to another investigator for future research studies without your additional informed consent.

The results of this research study may be presented in summary form at conferences, in presentations, reports to the sponsor, academic papers, and as part of a thesis/dissertation.

Can I be removed from the research without my OK?

The person in charge of the research study or the sponsor can remove you from the research study without your approval. Possible reasons for removal include technical issues.

What else do I need to know?

We will not offer to share your individual test results with you. You may accept or decline these results.

Signature Block for Capable Adult

Your signature documents your permission to take part in this research. We will provide you with a signed copy of this form for your records.

Signature of subject

Date

Printed name of subject

Signature of person obtaining consent

Date

Printed name of person obtaining consent

A.2 Pre-Experiment Survey

AR Brainstorming Pre-Experiment Study

Thank you for volunteering as a participant in our study! Please help us by completing the below survey before we start.

* Required

1. Name *

____

2. Gender

Mark only one oval.

○ Male
○ Female

3. What is your age? *

Mark only one oval.

○ 18-24 years old
○ 25-34 years old
○ 35-44 years old
○ 45-54 years old
○ 55-64 years old
○ 65-74 years old
○ 75 years or older

4. How familiar would you consider yourself to be with Augmented Reality headsets? *

Mark only one oval.

○ Extremely familiar
○ Very familiar
○ Moderately familiar
○ Slightly familiar
○ Not at all

5. How familiar would you consider yourself to be with the Microsoft HoloLens? *

Mark only one oval.

○ Extremely familiar
○ Very familiar
○ Moderately familiar
○ Slightly familiar
○ Not at all

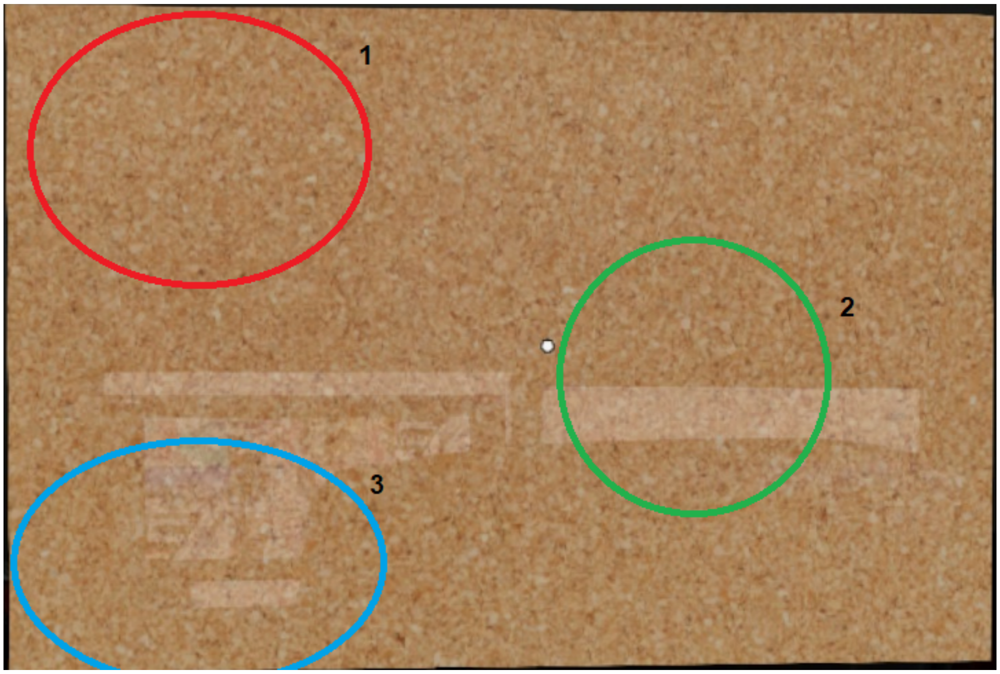
6. Have you ever participated in an affinity diagramming session before? *

Mark only one oval.

○ Yes
○ No

A.3 Category Drawing Template

Please approximate the locations and coverage of the categories you created with your partners, as shown in the following example. Number your categories starting at 1 so you can refer to them in a Google Forms survey.



A.4 Category Survey

Category Survey

* Required

1. Participant ID *

2. Session ID *

3. What is the name of the category indicated by the number 1?

4. What does the number 1 category mean to you?

5. What is the name of the category indicated by the number 2?

6. What does the number 2 category mean to you?

7. What is the name of the category indicated by the number 3?

8. What does the number 3 category mean to you?

9. What is the name of the category indicated by the number 4?

10. What does the number 4 category mean to you?

11. What is the name of the category indicated by the number 5?

12. What does the number 5 category mean to you?

13. What is the name of the category indicated by the number 6?

14. What does the number 6 category mean to you?

A.5 Post-Task Survey

Basic Information

* Required

1. Participant ID *

____

2. What was the date of your session (if you can remember)?

____

Example: January 7, 2019

3. What was the time of your session (if you can remember)?

____

Example: 8:30 AM

4. What was the cue condition that you just experienced? *

Mark only one oval.

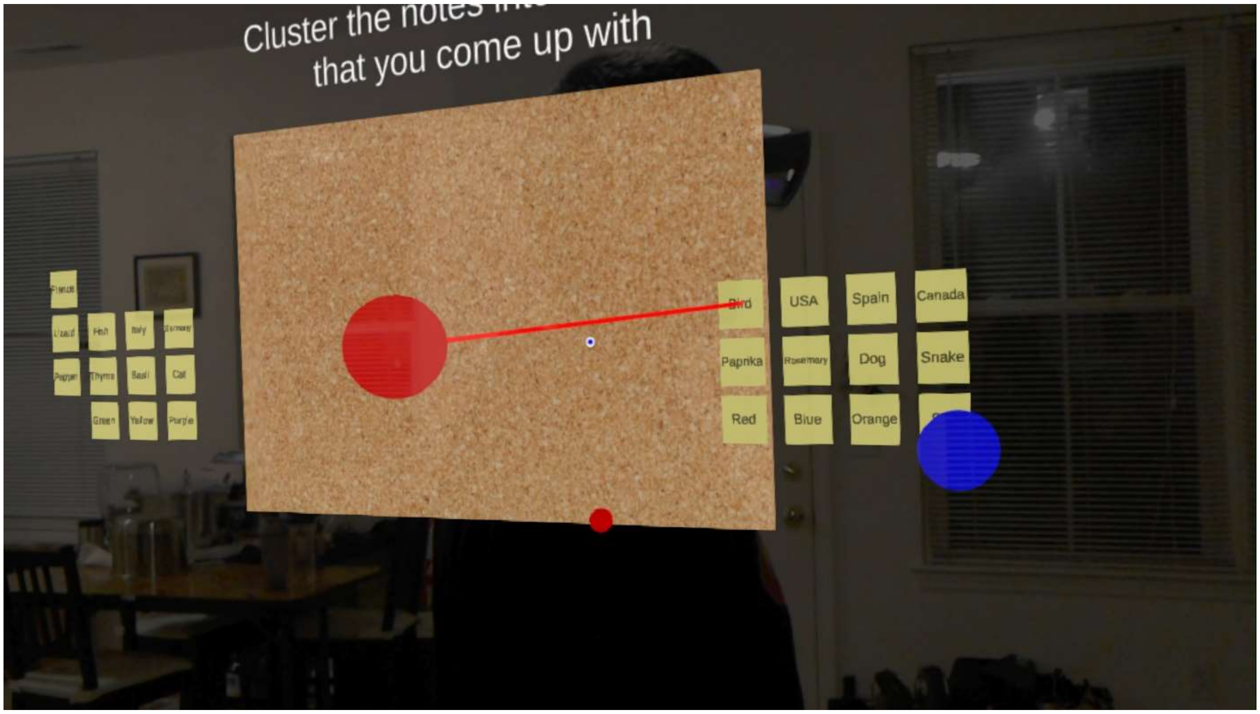
○ Cue-Full Condition (cues were visible) *Skip to question 5*

○ Cue-Less Condition (cues were NOT visible) *Skip to question 39*

Cue-Full Condition

Please answer the following questions specifically about the condition where collaborative cues were visible. As seen in the following image, this condition included visible rays indicating where other users were looking, simple avatars for user heads and hands, and outlines around notes when users focused them and when they moved them.

Screenshot of virtual environment condition WITH cues



5. I think that I would like to use this system frequently. *

Mark only one oval.

○ Strongly Disagree
○ Somewhat Disagree
○ Neutral
○ Somewhat Agree
○ Strongly Agree

6. I found the system unnecessarily complex. *

Mark only one oval.

○ Strongly Disagree
○ Somewhat Disagree
○ Neutral
○ Somewhat Agree
○ Strongly Agree

7. I thought the system was easy to use. *

Mark only one oval.

○ Strongly Disagree
○ Somewhat Disagree
○ Neutral
○ Somewhat Agree
○ Strongly Agree

8. I think that I would need the support of a technical person to be able to use this system. *

Mark only one oval.

○ Strongly Disagree
○ Somewhat Disagree
○ Neutral
○ Somewhat Agree
○ Strongly Agree

9. I found the various functions in this system were well integrated. *

Mark only one oval.

- ○ Strongly Disagree
- ○ Somewhat Disagree
- ○ Neutral
- ○ Somewhat Agree
- ○ Strongly Agree

10. I thought there was too much inconsistency in this system. *

Mark only one oval.

- ○ Strongly Disagree
- ○ Somewhat Disagree
- ○ Neutral
- ○ Somewhat Agree
- ○ Strongly Agree

11. I would imagine that most people would learn to use this system very quickly. *

Mark only one oval.

- ○ Strongly Disagree
- ○ Somewhat Disagree
- ○ Neutral
- ○ Somewhat Agree
- ○ Strongly Agree

12. I found the system very cumbersome to use. *

Mark only one oval.

○ Strongly Disagree
○ Somewhat Disagree
○ Neutral
○ Somewhat Agree
○ Strongly Agree

13. I felt very confident using the system. *

Mark only one oval.

○ Strongly Disagree
○ Somewhat Disagree
○ Neutral
○ Somewhat Agree
○ Strongly Agree

14. I needed to learn a lot of things before I could get going with this system. *

Mark only one oval.

○ Strongly Disagree
○ Somewhat Disagree
○ Neutral
○ Somewhat Agree
○ Strongly Agree

15. How do you think you performed on the task?

16. What was the most challenging part of using the tool?

17. What was the most challenging in collaboration and communicating with other people, especially regarding referencing virtual objects?

18. When you worked, to what extent did you pay attention to what other people are doing?

19. Was it ever difficult to draw attention to specific objects?

20. Do you feel the visible cues (including gaze rays, highlights) helped with the task?

21. Was having cues at all distracting to the collaboration or focusing on your own tasks?

22. Do you have ideas for tools that might have made your task easier?

Cue-Less Condition

Please answer the following questions specifically about the condition where collaborative cues were NOT visible. As seen in the following image, this condition did not include any visual feedback for what other users were doing or when a note was being moved.

Screenshot of virtual environment WITHOUT cues

23. I think that I would like to use this system frequently. *

Mark only one oval.

○ Strongly Disagree
○ Somewhat Disagree
○ Neutral
○ Somewhat Agree
○ Strongly Agree

24. I found the system unnecessarily complex. *

Mark only one oval.

- ○ Strongly Disagree
- ○ Somewhat Disagree
- ○ Neutral
- ○ Somewhat Agree
- ○ Strongly Agree

25. I thought the system was easy to use. *

Mark only one oval.

- ○ Strongly Disagree
- ○ Somewhat Disagree
- ○ Neutral
- ○ Somewhat Agree
- ○ Strongly Agree

26. I think that I would need the support of a technical person to be able to use this system. *

Mark only one oval.

- ○ Strongly Disagree
- ○ Somewhat Disagree
- ○ Neutral
- ○ Somewhat Agree
- ○ Strongly Agree

27. I found the various functions in this system were well integrated. *

Mark only one oval.

- ○ Strongly Disagree
- ○ Somewhat Disagree
- ○ Neutral
- ○ Somewhat Agree
- ○ Strongly Agree

28. I thought there was too much inconsistency in this system. *

Mark only one oval.

- ○ Strongly Disagree
- ○ Somewhat Disagree
- ○ Neutral
- ○ Somewhat Agree
- ○ Strongly Agree

29. I would imagine that most people would learn to use this system very quickly. *

Mark only one oval.

- ○ Strongly Disagree
- ○ Somewhat Disagree
- ○ Neutral
- ○ Somewhat Agree
- ○ Strongly Agree

30. I found the system very cumbersome to use. *

Mark only one oval.

○ Strongly Disagree
○ Somewhat Disagree
○ Neutral
○ Somewhat Agree
○ Strongly Agree

31. I felt very confident using the system. *

Mark only one oval.

○ Strongly Disagree
○ Somewhat Disagree
○ Neutral
○ Somewhat Agree
○ Strongly Agree

32. I needed to learn a lot of things before I could get going with this system. *

Mark only one oval.

○ Strongly Disagree
○ Somewhat Disagree
○ Neutral
○ Somewhat Agree
○ Strongly Agree

33. How do you think you performed on the task?

34. What was the most challenging part of using the tool?

35. What was the most challenging in collaboration and communicating with other people, especially regarding referencing virtual objects?

36. When you worked, to what extent did you pay attention to what other people are doing?

37. Was it ever difficult to draw attention to specific objects?

38. Do you have ideas for tools that might have made your task easier?

Cue-Less Condition

Please answer the following questions specifically about the condition where collaborative cues were NOT visible. As seen in the following image, this condition did not include any visual feedback for what other users were doing or when a note was being moved.

A.6 External Judges Consent Form

Title of research study: *IRB 21-170 Evaluating Collaborative Cues in AR Affinity Diagramming Tasks*

Principal Investigator: *Sang Won Lee* *sangwonlee@vt.edu* *(540)231-4857*

Other study contact(s):

- *Doug A. Bowman* *dbowman@vt.edu* *(540) 231-2058*
- *Nathaniel Llorens* *lnathaniel@vt.edu* *(561) 400-4804*

Key Information: The following is a short summary of this study to help you decide whether or not to be a part of this study. More detailed information is listed later on in this form.

We invite you to participate in this Augmented Reality (AR) study to help us evaluate the results of previous study sessions. This work is hoping to make AR more usable and we deeply appreciate your participation.

If you agree to participate, you will use a HoloLens headset to view the start and end configurations of experiment sessions and will then rate them based on organization and understandability via an electronic survey. The total time commitment for this study is around 100 minutes and you will be compensated with a $18 Amazon gift card for your time.

Detailed Information: The following is more detailed information about this study in addition to the information listed above.

Who can I talk to?

If you have questions, concerns, or complaints, or think the research has hurt you, talk to the research team at *lnathaniel@vt.edu / (561) 400-4804.*

This research has been reviewed and approved by the Virginia Tech Institutional Review Board (IRB). You may communicate with them at 540-231-3732 or irb@vt.edu if:

- You have questions about your rights as a research subject
- Your questions, concerns, or complaints are not being answered by the research team
- You cannot reach the research team
- You want to talk to someone besides the research team to provide feedback about this research

How many people will be studied?

We plan to include about 2-4 people in this research study.

What happens if I say yes, I want to be in this research?

The study will take place in the Usability Lab 133 (D, E, F) (102 McBryde Hall) and will take approximately 100 minutes. When you arrive, you will be greeted and asked to read and sign the informed consent form after your questions (if any) are answered. Then you will be asked to clarify you have normal vision. Next you will be provided with written or verbal instructions for the experiment, and familiarized with the lab and the equipment they will be using. You will wear an augmented reality (AR) headset such as the Microsoft Hololens. Using the device, you will then be shown start and end views from previous sessions, along with notes given by the participants, and answer questions in an online survey to rate previous performance. Tasks will involve physical movements including looking around the environment. You will be presented with 4 of these start and end views and given approximately 20 minutes to rate each pair. The entire session will be audio recorded.

To thank you for your participation, we will give out a $18 electronic gift card at the end of the study.

What happens if I say yes, but I change my mind later?

You can leave the research at any time, for any reason, and it will not be held against you.

If you decide to leave the research, contact the investigator so that the investigator can terminate the current experiment session. Any data collected by far will be disregarded and will not be used for further study. There will be no follow-up procedures afterwards.

Withdrawing from the experiment does not affect receiving $12 gift card compensation.

Is there any way being in this study could be bad for me? (Detailed Risks)

Using AR technology can produce symptoms of sickness or discomfort in some users. These symptoms are usually mild, and may include dizziness, nausea, eye strain, headache, or disorientation. During tasks involving physical movement, there is also some risk that you could collide with obstacles in the physical environment.

You will be given the option to take a break or quit the experiment at any time. To mitigate the risk of sickness and discomfort, we will adjust the display properly, keep task sessions short, provide frequent breaks, and ask you after each set of tasks how you are feeling. To mitigate the risk of physical obstacles and cabling, we will clear the area of obstacles, show you where the boundaries of the space are.

What happens to the information collected for the research?

We will make every effort to limit the use and disclosure of your personal information, including research study, only to people who have a need to review this information. We cannot promise complete confidentiality. Organizations that may inspect and copy your information include the IRB, Human Research Protection Program, and other authorized representatives of Virginia Tech.

All data collected during this study will be done so anonymously and stored in a password protected computer. No names, contact information, or any other identifying information will be attached to your responses to an investigator’s questions or to your results from the provided tasks. At no time will the researchers release identifiable results of the study to anyone other than individuals working on the project without your written consent. However, you may be asked to perform tasks in a public space and bystanders could recognize you.

If identifiers are removed from your private information or samples that are collected during this research, that information or those samples could be used for future research studies or distributed to another investigator for future research studies without your additional informed consent.

The results of this research study may be presented in summary form at conferences, in presentations, reports to the sponsor, academic papers, and as part of a thesis/dissertation.

Can I be removed from the research without my OK?

The person in charge of the research study or the sponsor can remove you from the research study without your approval. Possible reasons for removal include technical issues.

What else do I need to know?

We will not offer to share your individual test results with you. You may accept or decline these results.

Signature Block for Capable Adult

Your signature documents your permission to take part in this research. We will provide you with a signed copy of this form for your records.

Signature of subject	Date
Printed name of subject	
Signature of person obtaining consent	Date
Printed name of person obtaining consent	

A.7 External Judges Survey

External Judges Survey

* Required

1. Name *

____

2. Session number *

____

3. The participants had broad agreement on the categories being used.

Mark only one oval.

○ Strongly Disagree
○ Disagree
○ Somewhat Disagree
○ Neutral
○ Somewhat Agree
○ Agree
○ Strongly Agree

4. For most notes, it was easy to understand why they were placed in a category.

Mark only one oval.

○ Strongly Disagree
○ Disagree
○ Somewhat Disagree
○ Neutral
○ Somewhat Agree
○ Agree
○ Strongly Agree

5. The proposed categories are clearly defined such that it is easy to tell which notes belongs to which categories.

Mark only one oval.

○ Strongly Disagree
○ Disagree
○ Somewhat Disagree
○ Neutral
○ Somewhat Agree
○ Agree
○ Strongly Agree

6. Ideas related to each other are grouped together in the same category.

Mark only one oval.

○ Strongly Disagree
○ Disagree
○ Somewhat Disagree
○ Neutral
○ Somewhat Agree
○ Agree
○ Strongly Agree

7. It is easy to tell how many categories there exist at a glance.

Mark only one oval.

○ Strongly Disagree
○ Disagree
○ Somewhat Disagree
○ Neutral
○ Somewhat Agree
○ Agree
○ Strongly Agree

8. The participants utilized the entire virtual space effectively (e.g., visually balanced, decluttered, no overlap).

Mark only one oval.

○ Strongly Disagree
○ Disagree
○ Somewhat Disagree
○ Neutral
○ Somewhat Agree
○ Agree
○ Strongly Agree