Using Mixed Reality to Support Face-to-Face Social Interactions

SMALL GROUP STUDY

05688 A – Adv. John Zimmerman
Carnegie Mellon University
December 12, 2017

Aaron Faucher
Grace Guo
Ravi Morbia
Frank Teng

Introduction

The outset of our project began at the intersection of two key questions.

1. How do people respond to AR in a social context?

2. How might AR improve the experience of a professional networking event?

1 How do people respond to AR in a social context?

Much has been written about the social implications of some of the first headworn computing devices – namely, Google Glass – and the social faux pas of this device. Non-users felt uncomfortable around the Glass, and were never sure if they were being recorded; wearers were dismissed as elitist technophiles disconnected from reality. One could be forgiven for assuming that Google Glass's initial failures would negate AR's chances of becoming a commonplace technology in the future.

Fast forward to 2017, and the landscape of augmented reality couldn't appear further from this fate. Magic Leap, a secretive mixed reality startup in Florida, is valued at north of $6 Billion. Mixed reality sits at equal footing with artificial intelligence in Microsoft's corporate strategy. Apple is stealthily building their own AR capabilities through numerous acquisitions and their release of ARKit. Snapchat's Spectacles, while unsuccessful in the market, hardly faced the same backlash as Glass; rather, they were briefly a trending mark of fashion. Increasingly, it appears likely that headworn mixed reality is a technology that future generations may contend with.

Our team sought to build off an assumption: if mixed reality is to become a prevalent technology in the coming decades, what types of social and cognitive effects can we expect from ubiquitous information displays? Social AR displays have been postulated in several concept films, but most utopian and dystopian futures depicted by VFX artists have not evaluated the real and felt effects of these devices on user cognition and social-emotional experience. Our team hoped to build a functional prototype that might allow us to explore these questions.

2

How might AR improve the experience of a professional networking event?

In order to narrow into a specific social context to explore the social effects of AR, our team felt it necessary to focus on a specific domain. Some of the early domains discussed included an exploration of social interactions between service workers and customers in a retail environment – how might AR improve this interaction by augmenting the conversation with relevant data from prior service interactions? We also briefly considered the possibility of doctor-patient interactions, where a doctor has a high need to focus attention on the patient and maintain a socially comfortable conversation, while also referencing supplementary information on the patient's medical history.

Ultimately, we were prompted by our faculty advisor, John Zimmerman, to consider the domain of professional networking. What sort of latent value is lost at networking events today? How are common human biases, such as homophily, reducing the efficacy of professional networking sessions meant to form new connections? How might we create more meaningful and effective networking experiences that connect the dots, push us outside of our comfort zones, or create new comfort zones? Professional networking, then, became the focus for us to explore these questions about AR and social interaction.

Discover

Define

Develop

Deploy

Observation
Fly-on-the-Wall

Conducting Observations

Our discovery phase explored these two questions in sequence. Rooted in the domain of professional networking, we first used both fly-on-the-wall and participatory observation methods to expose ourselves to the needs and motivations inherent in the networking context. We split into two groups and each attended a local Pittsburgh networking session: 'Build Night' at Pittsburgh Code & Supply, and 'Machine Learning for Social Good' at CMU's Rangos Ballroom.

Pittsburgh Code & Supply Metup

Pittsburgh CODE & SUPPLY Co.


Rangos Ballroom, CMU
Machine Learning for Social Good Meetup

Observation Findings

Points of Connection

We discovered that points of connection were excellent conversation starters, and helped break the ice in a new relationship. For example, at one networking event, a student wearing a University of Pittsburgh sweatshirt was approached by an individual asking more specific questions about their major, their classes, and ultimately whether they knew a specific professor – presumably because they were also an alumnus. An AR solution that suggested possible points of connection, in the same way that a Pitt sweatshirt suggested possible points of connection, might lead to a positive user experience.

Facilitators

We observed facilitators playing an important role in making connections between people who were new, or might not otherwise meet. At Code & Supply, the facilitator started the meetup by prompting brief introductions by all in attendance. For a student (first-timer) who was attempting to learn a new coding language, the facilitator connected him with a returning member who he knew had experience in this language. This connection may have happened by chance, but the facilitator greatly accelerated the connection and improved the experience for the student. An AR solution that might facilitate these otherwise hidden connections would leverage this opportunity..

Structure

We also observed the discomfort that can follow a lack of structure at a networking event. For first-timers who may not be familiar with social norms at the event, a lack of order or direction can cause distress on top of existing anxieties associated with meeting new people. At Pittsburgh Code & Supply, we experienced this ourselves, as we sat quietly around tables waiting for the event to start, unsure whether we should be getting food, starting conversation, or waiting for a more structured beginning. As such, we felt that a possible area for exploration was an AR solution that might add more structure to the process of socializing, and reduce the feeling that one might be breaking social norms.

Objects

We saw that objects played a promising role – having something to talk about besides either person – in starting conversations. At another event, unique knick-knacks set on a table helped prompt new conversations. As such, this prompted the idea that an AR solution might be effective if it provided things to talk about.

Exploratory Bodystorming

Following our observations, the team embarked on an exploratory phase of research through design. Building on our initial findings, we ran several bodystorming sessions in which we quickly exposed fundamental questions about integrating AR into a social setting. Using these findings, we rapidly generated paper prototypes and started play-acting scenarios with students outside our group to test design hypotheses.

Our internal bodystorming sessions raised more questions than they answered, but helped to direct our later prototyping and ideation as we honed in on answers. Some of the key questions raised by bodystorming with paper prototypes included the following.

? **Should an AR experience be active or passive?**
Does the user control the AR overlays using clicks, taps, or voice, and control when/how information appears in the display, or do overlays appear intelligently so the user can focus on the conversation?

? **How might we avoid uncomfortable feelings of self-consciousness about one's own AR overlay?**
Does the user know what the other person can see about them?

? **How much data, and what data, should be shown?**

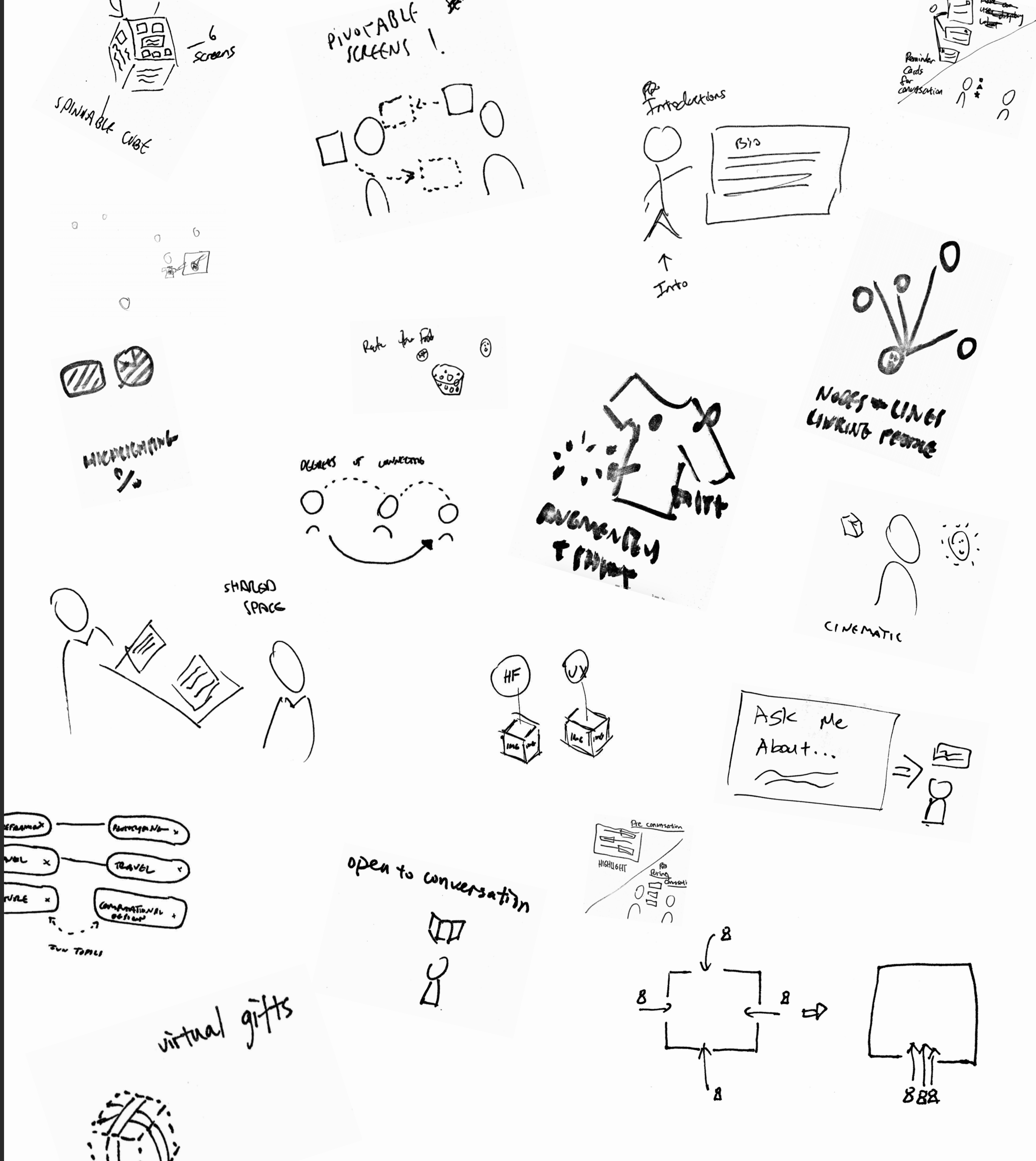
? **Does it feel more natural to provide detailed information on someone, or prompts for conversation?**

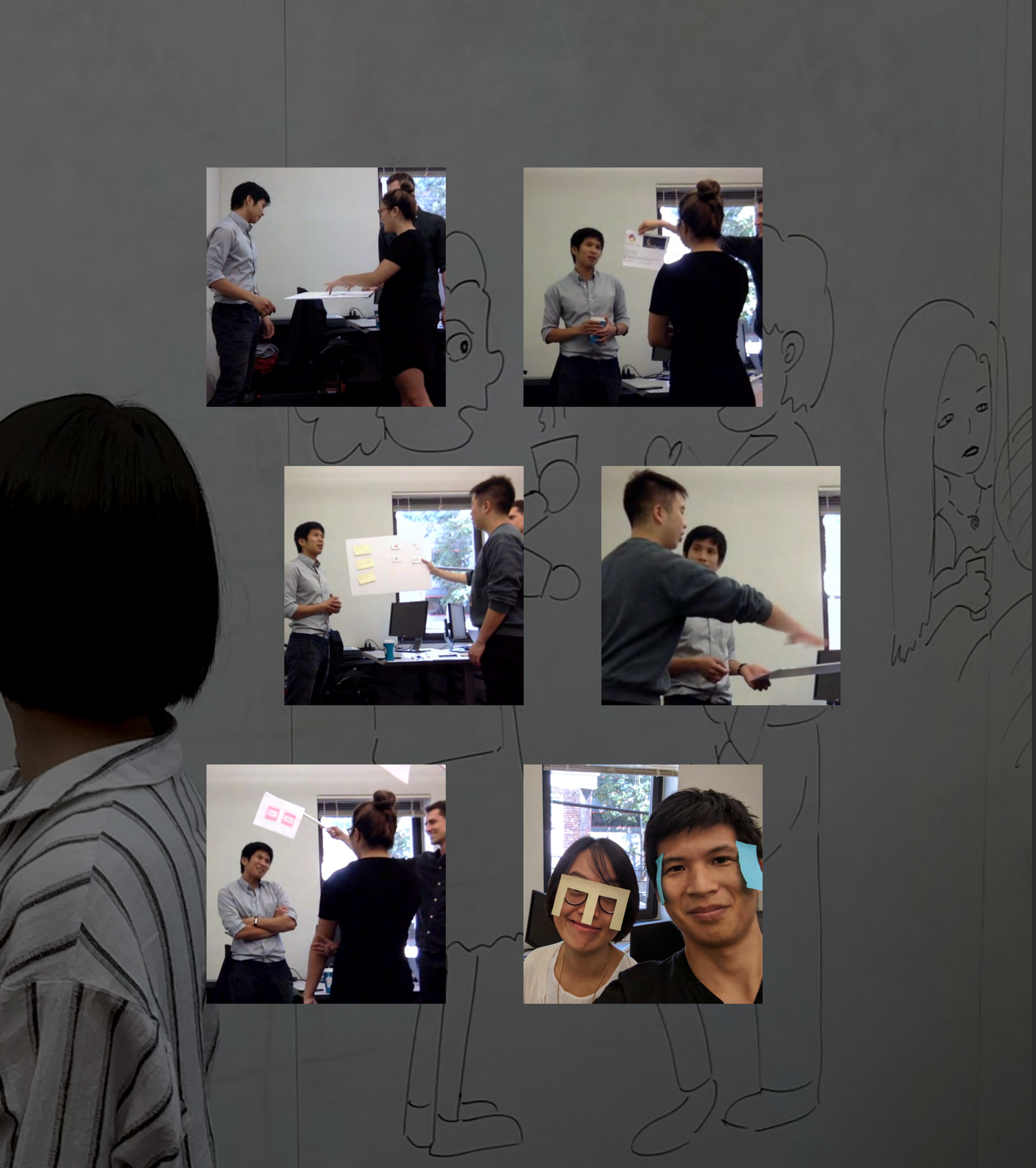
? **How might we make a high volume or breadth of information about a person easily digestible in the high-cognitive-load context of a conversation?**

? **How do we communicate that a conversation partner is still paying attention, when they may be reading an AR overlay?**

Ideation

Based on initial findings, our team did several rounds of paper-pen sketching and ideation. Generative methods included Google Ventures' Crazy Eights format, 'Yes And' improv, and Round Robin. This generative sprint helped to expand our concept of the possibilities when moving towards the next step of paper prototypes.





User Tests
Low-Fidelity

Following bodystorming and ideation, our team rapidly generated low-fidelity prototypes with foamcore, paper, and transparencies, meant to test along a few parameters:

- user preference for active vs. passive AR systems
- user preference for information vs. conversation prompts
- user preference for private AR space vs. shared AR space

We were able to conduct two user tests on these low fidelity prototypes. We ran these tests using another graduate student as a participant, a researcher as a conversation partner, and another researcher manipulating the low-fidelity prototypes.

These low-fidelity tests led to a number of findings. First, we noticed that the spatial placement of AR information had a significant effect on user's social comfort. For example, one user felt quite uncomfortable when they didn't know what information was hovering above their head, and kept checking it to see. In another test, when we used a shared, interactive AR board to communicate information about both parties, one user described feeling snubbed when the other user started interacting with the device more than talking with him.

Discover

Define

Develop

Deploy

Investigating Questions

Armed with several findings from these early, exploratory methods, the team moved on to three simultaneous prototyping threads meant to investigate these individual questions more in-depth, and converge on an ideal solution to build into a functional AR prototype.

Data & Privacy

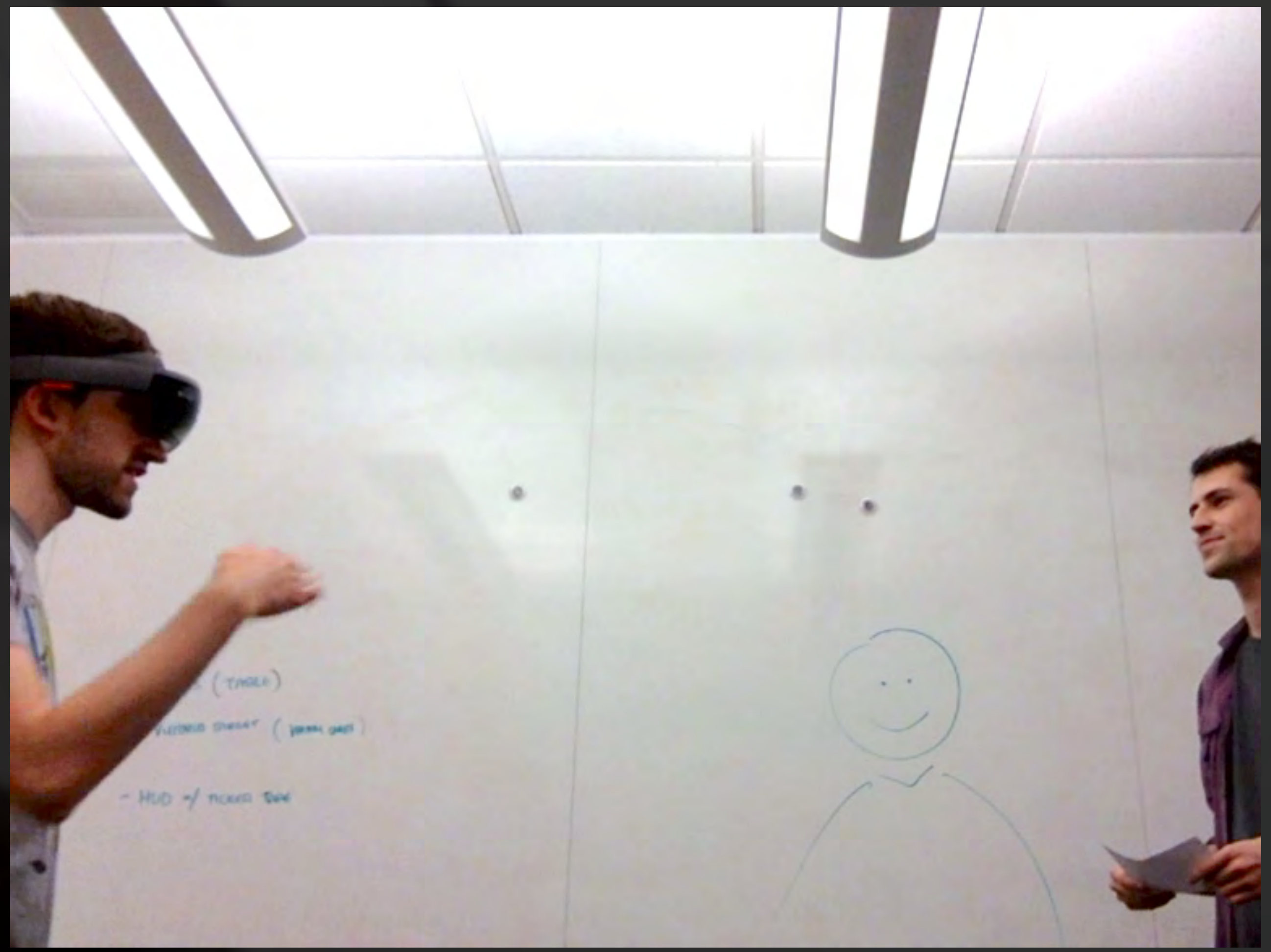
One group began speed-dating several storyboard permutations concepts specifically to understand users' feelings around data and privacy: how much control would they want over their data? Do they have a preference for a shared (i.e. viewable by both parties) AR experience, or private experience (i.e. my conversation partner is restricting from viewing what I am seeing about them)? This team's general take-away was that users had similar feelings about AR as they did about today's digital landscape – namely, that if information was publicly available via Google, or if they had offered that information specifically, then they were comfortable with it appearing on their overlay, and with other users browsing that information in a networking session.

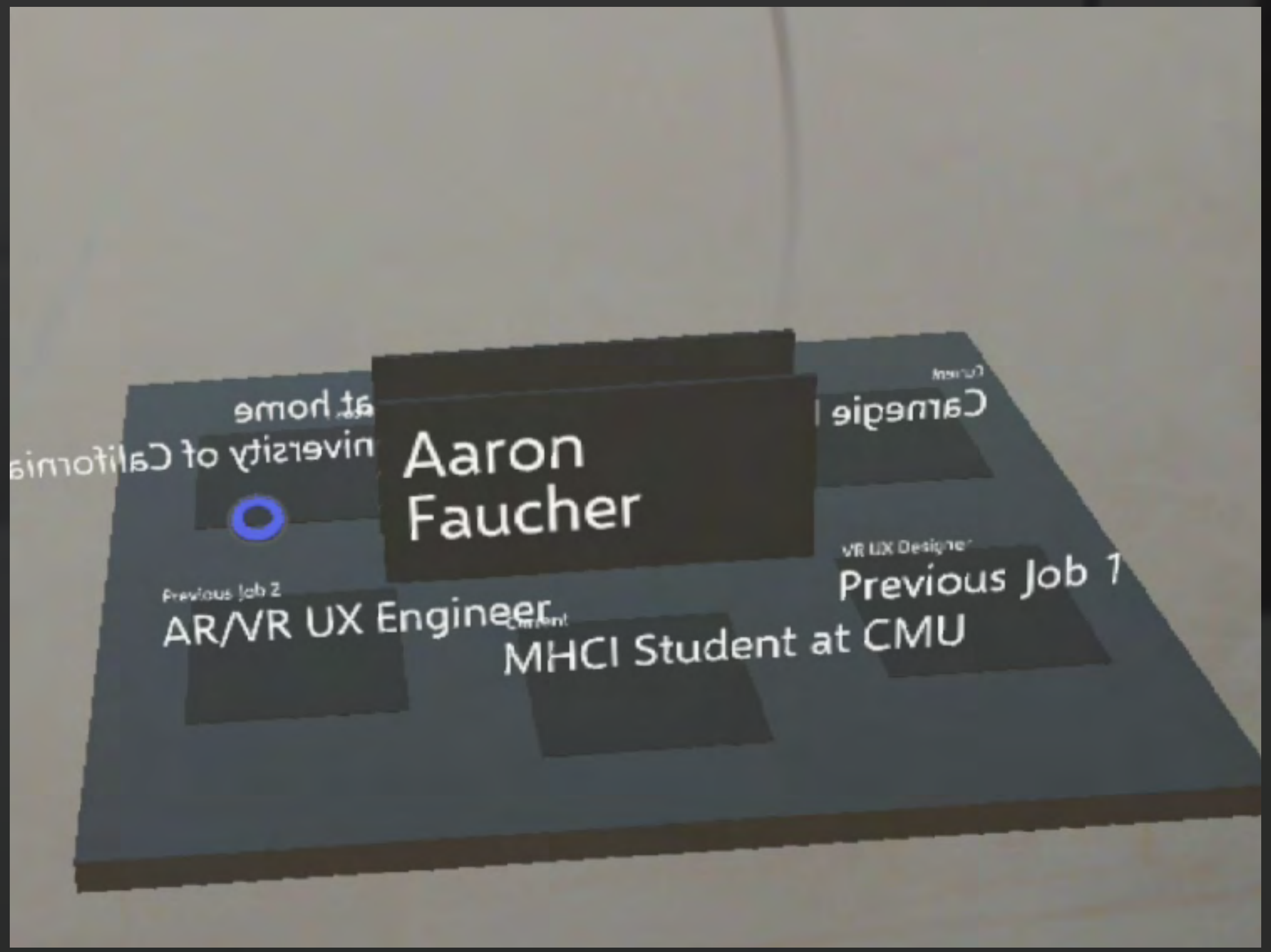
Sharing & Placement

Another group began speed-dating several new low-fidelity prototypes meant to better understand the relationship between interactivity, active vs. passive AR, and the impact of spatial placement on conversation. They managed to converge on a hybrid interface that allowed for private browsing and interactivity with AR components, but a sharing feature that would then allow this content to be placed in shared space for discussion.

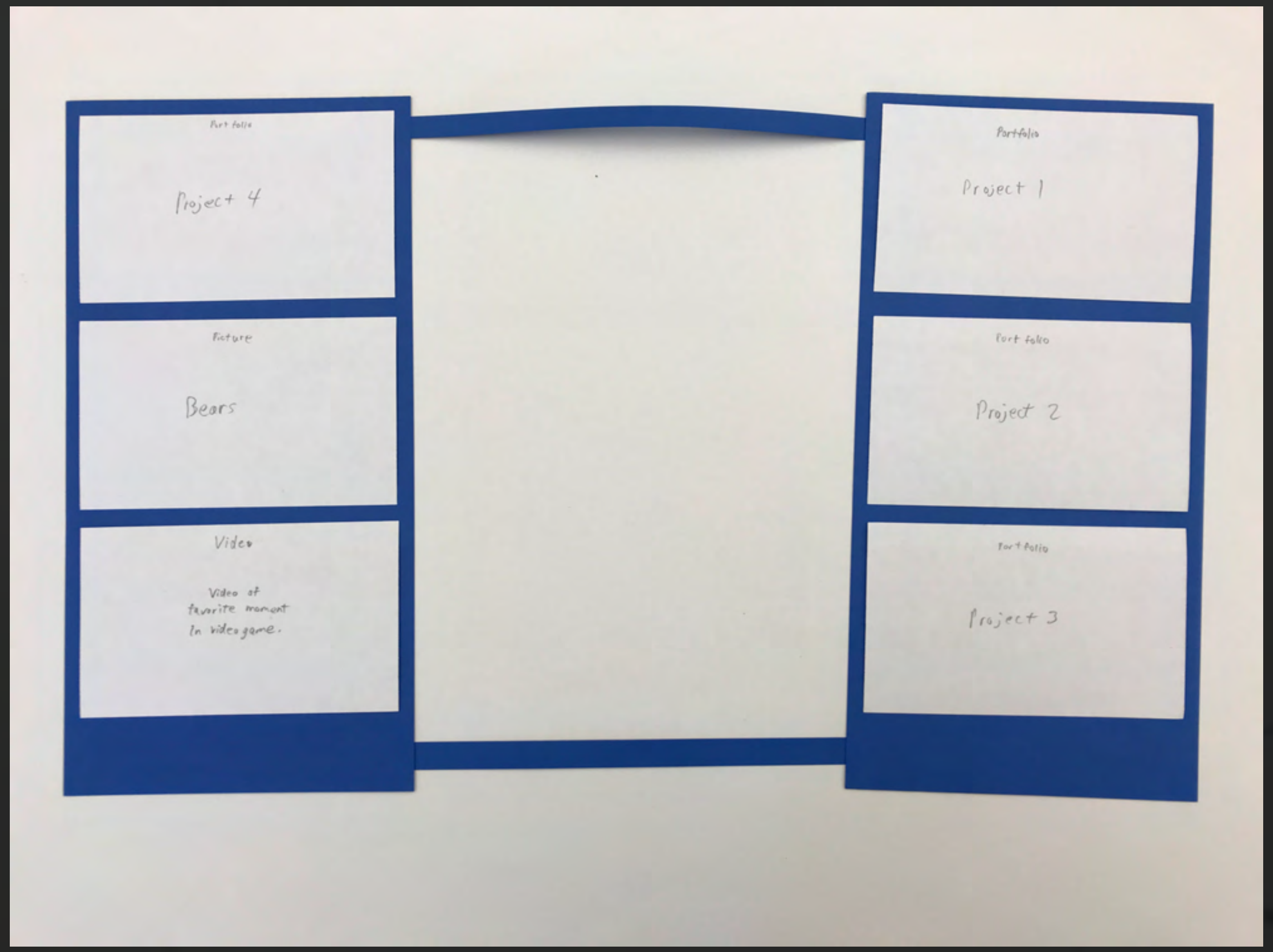
AR Factors

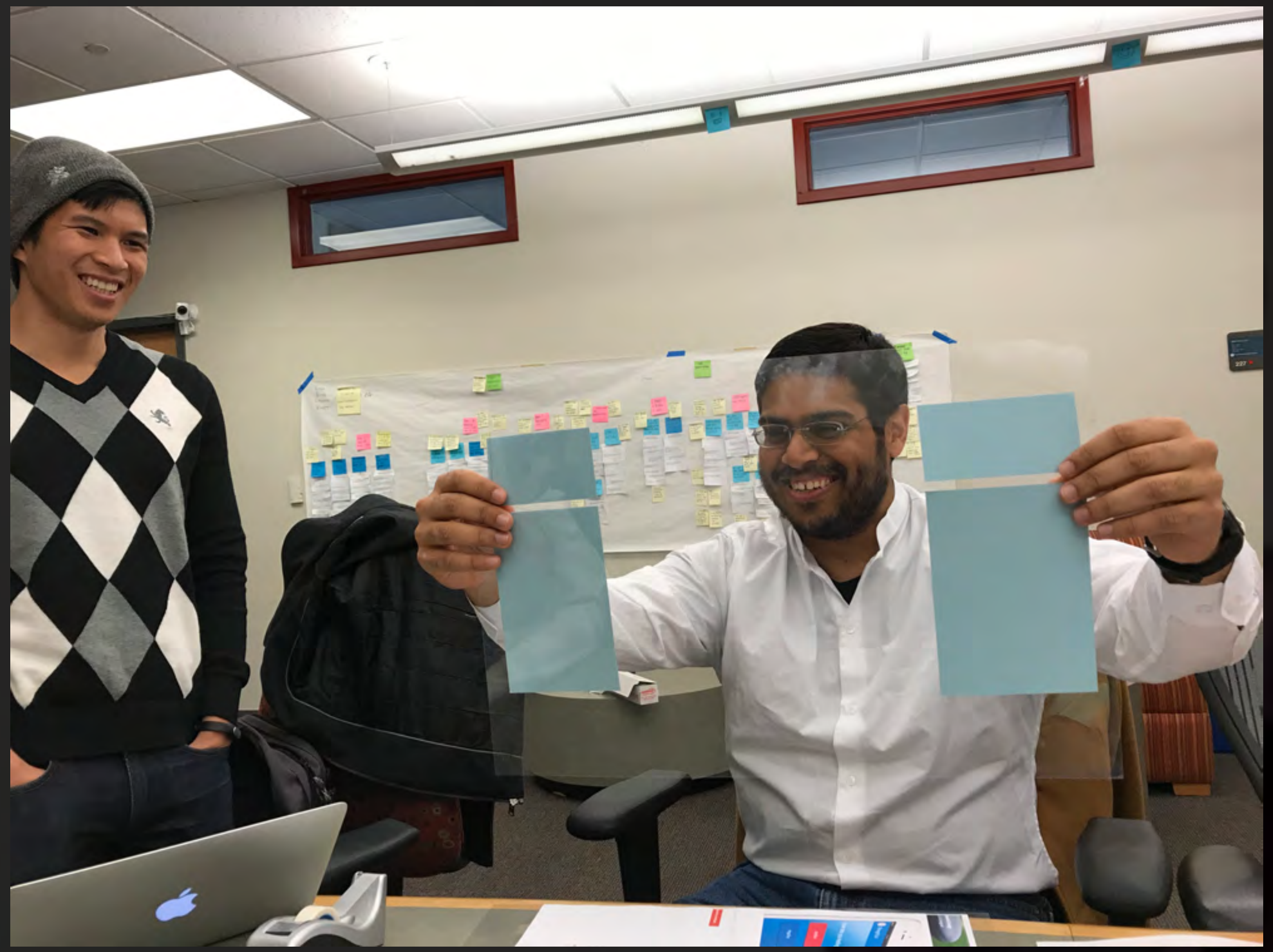
Yet another team concurrently prototyped two mid-fidelity AR experiences on the Microsoft HoloLens, and tested with two more users in a conversational setting. These tests led to valuable findings around the effects of certain technological constraints of AR: namely, the narrow field of view obscured certain interface elements, which led to discomfort during conversation as the user craned their neck to view information. It also highlighted limitations of user cognition in the conversational setting – both users expressed discomfort in maintaining a comfortable pace of conversation while scanning sentences of text next to the user's head. It also helped confirm that interface elements around the conversation partner, rather than close to the user, were more comfortable to interact with. Fundamentally, users in this group described a desire to know information prior to conversation, rather than at the moment of contact.





Aaron Faucher
Previous Job 2
AR/VR UX Engineer
MHCI Student at CMU
Previous Job 1





Early
Findings

Pursuing these three separate threads helped expose initial findings that would be used to define the final concept for production in AR.

People want information before the conversation.

Some of our early testing yielded comments from users that they would prefer to see info about a person before approaching him or her, rather than during the conversation. Seeing the info beforehand lowered anxiety and informed users' decisions regarding whom to approach. Minimizing information shown once the conversation began also reduced the cognitive load on testers.

People are okay with sharing information under the right conditions.

Much of our initial research focused on how users might use AR during conversations. We rapidly ideated dozens of ideas that ranged from a UI that guides the user around the room with points of interest, to sharing files, images, etc. while in a conversation (Ideation Scans). To validate or disconfirm our assumptions, we created storyboards and speed dated several people. There were three parameters we tested: private vs. shared screens, manual vs. automated AR, and the level of privacy of information. The results showed that testers preferred having their interface be private by default, but with the option to share it. Users also indicated that they do not mind if others look up information about them online without explicit permission, since this is doable on any other internet-enabled device. Furthermore, having to give permission every time becomes annoying, and saying "no" would seem suspicious.

People don't want to read sentences or paragraphs of text.

In our testing, we also found that users had difficulty processing even one paragraph of text so we reduced the information shown about others to short terms and phrases. Additionally, we experimented with the idea of using images rather than text to act as a visual aid to the conversation. In a networking context, this might be a portfolio piece or an image related to another interest.

In AR, this ability would be particularly impactful for those who have work involving 3D models (e.g. architects, game designers, motion graphic artists, etc.). We tested how this might look with users, employing low-fi paper prototypes [photos of shared screen prototypes]. We tested the placement of shared objects between users and off to the side, and found that users prefer to have the shared object directly between them and their partner in conversation (as long as the object is partially transparent).

Refocus on pre-conversation interaction.

Though sharing images of work in AR would be useful for many professions, it would not necessarily benefit the majority, who would only have 2D images to share, something that is already possible with phones and tablets. As such, we shifted focus to tackle a problem that has not yet been solved: breaking homophily. It is natural for most people to gravitate towards like themselves, and without any deeper information, this means users tend to speak to others of the same gender, race, age. In order to overcome this inclination, we designed a tag system that would show potential areas of overlapping interest to both parties.

Narrowing Concept

As our AR prototype progressed, we narrowed down to a core feature set, which included a tagging system with filtering. Each user would be able to input short terms or phrases about themselves for others to see. The interface then allows each person to select one or more tags to filter the display such that others' tags that match the user's selection are highlighted. This enables users to quickly scan a room with dozens of people and easily find every person with a matching interest.

We settled on a TAGGING concept because…

1. users wanted info before talking
2. tags are quick to scan
3. tags are a familiar metaphor
4. tags are atomic units that can be matched

Refining Concept

Using index cards and sticky notes to represent tags, we gathered a group of graduate students and quickly tested the paper-versioned tagging system. We asked compare the current networking status quo situation with the rough tagging prototype. Our findings summarized as follows:

Users would prioritize who they would talk to based on tags about goal/interest.

User felt that tags added relevance, but they also asked for categories to understand whether someone had a certain tag as 'expertise' or interested.

Users wanted names and job titles in addition to tags.

Using locations or companies as tags confused users.

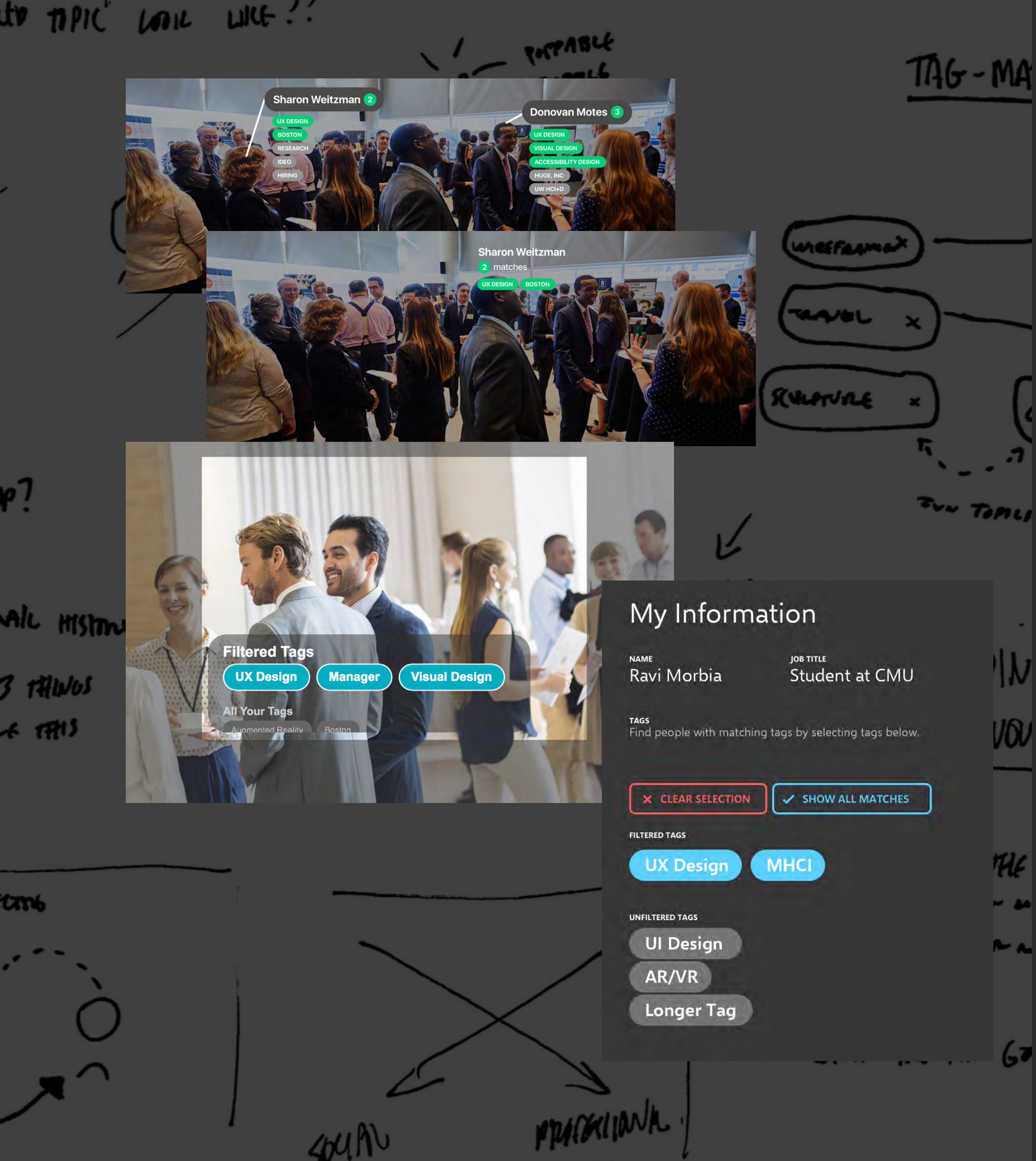
Users still wanted conversation starters and to see tags that didn't match theirs for the purpose of chatting with the person and asking about themselves.

This user testing helped to affirm a tagging system that displays tags entered by users, and allows users to highlight matching tags with others. Based on the previous findings, we envisioned our tagging system would feature these capabilities:

Allow user to enter tags that describe their background, which should be done before a networking event takes place.

Display the other users' name, job title, and entered tags next to their faces when they enter the primary user's line of sight.

Highlight some or all matching tags between primary user and other users.



UI Design

When developing the specifics of the interface, we looked at web UI conventions as well as conversational interfaces in video games. In order to keep the cognitive load minimal for users, we finalized our interface with just a few elements: the user's name, job title, and up to 5 tags describing their work and interests. The final UI also included two buttons, one of which allowed users to select all their tags to be matched, and the other that cleared all filters.

Discover Define **Develop** Deploy

Planning
Development

With the pieces collected, we began prototyping for HoloLens with Unity3D Engine and MRTK Toolkit. We split all development tasks into five workflows. Each workflow had low dependency on the others, allowing easier divide-and-conquer. The five workflows were:

Dynamic app data via Google Forms-Unity integration
User state management
HoloLens target tracking
Tag rendering & filtering functionality
Building user interface

An early functional prototype of the tag-filtering system, as viewed through the HoloLens.

Dynamic User Input

While developing the interface, we also explored a few solutions to the problem of how to integrate realtime user data (name, job title, and tags) and parse the data for use in the tagging system. An ideal way to get user data is a user-friendly front-end page for users to fill in their information. After comparing alternatives, we decided to use Google Form for its simple-to-use interface, plus the fact that the responses collected could be easily formatted as JSON.

User inputs information
and tags via Google Forms.

Google Forms outputs to a
Google Sheet.

Unity uses REST to access
sheet data, parses JSON,
and integrates user data into AR scene.

Target Tracking

We soon encountered a few technical challenges due to the unstable development environment for HoloLens and Augmented Reality in general. The most prominent challenge came from target tracking functionality with Vuforia. On the recently released Unity 2017.2, using Vuforia caused loss of interactivity on HoloLens which means that the tagging system wouldn't work if we wanted target tracking with Vuforia. Since the more important question we wanted to test out with the prototype didn't require target tracking to work in its perfect state, we decided to compromise on target tracking ability. Our workaround was to use fixed world anchors to place tags and then have users stay at the location where the tags are set.

This image was captured through our HoloLens – we were able to link augmented objects to real-world, dynamically changing positions using fiducials and the Vuforia computer vision toolkit. Vuforia demo scene.

Finalizing Prototype

Over the span of 2 days, we finished building all the features and were ready to test on HoloLens. Testing within the team, we uncovered a few usability issues and were able to fix most of them:

Added user training section specifically for user test since most users are not familiar with the gestures on Hololens.

Floating filter panel: not to exceed waist height, prevent the panel covering up tags on other participants.

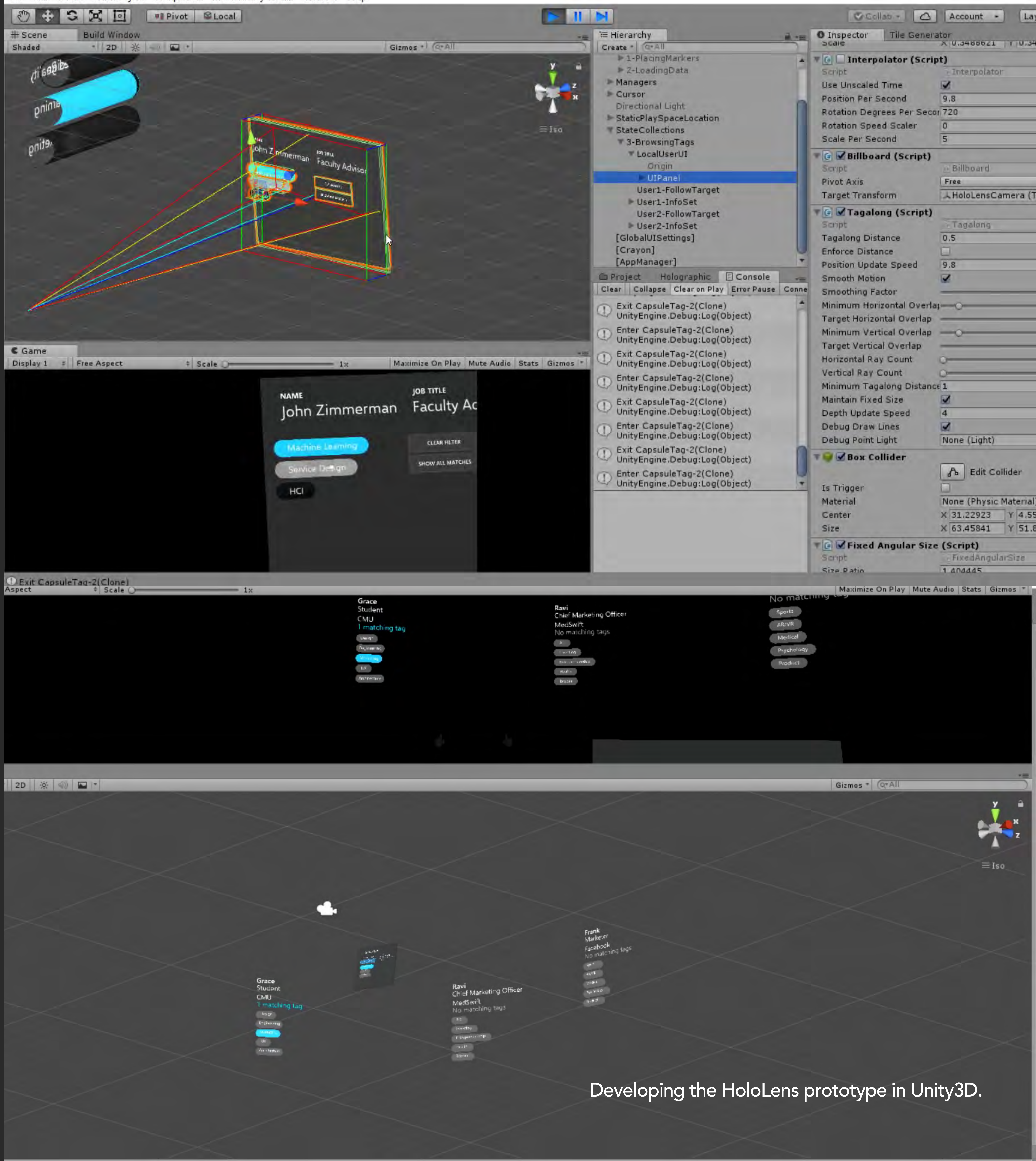
Visibility of text: dark colors were not rendered effectively on the HoloLens

In summary, here are the final set of features for the working prototype for HoloLens:

World anchor

Gestural training for new users

Dynamic data integration into 3D scene via Google Forms
Tagging system that:
- displays name, job title, organization, and tags entered
- highlights all matching tags between local user and other users
- allows filtering and highlighting tags on other users based on own tags



Developing the HoloLens prototype in Unity3D.

Discover Define Develop Deploy


Grace Guo
Student
Carnegie Mellon
1 matching tag
Medical
Engineering
Design
Ravi Morbia
CTO
MedSwift, Inc.
2 matching tags
Medical
Wearables
Analytics
Frank Teng
Partner
Pulse Ventures
AI
Investing
Entrepreneurship
Filter by Tag
Find people with matching tags by selecting tags below.
SHOW ALL MATCHES
CLEAR SELECTION
TAGS
Medical
Wearables
Marketing
Smartwatches
Health
MY INFO
Aaron Faucher
Marketing Coordinator
EKG, Inc.

Study Design

Our study aimed to test the 'Approach' stage of a networking event through a tag filtering system. Due to the resource constraints of modern AR technology, we could not test the full extent of the envisioned networking event where everyone wore AR-enabled headsets. In lieu of this, we focused on how might a user prioritize potential conversation partners in a room of three confederates. We recruited 5 participants from various industry backgrounds, with the requirement that they already had previous networking experience. Our participants were all Master's students between the ages of 25 - 35.

Introduce Persona

Our study asked our participants to imagine themselves as a fictitious persona. The persona constrained our participants to a controlled environment from which to test variables, and encouraged our subjects to step into an imaginary future. To add to the ambience, we played background crowd noises in a large room to simulate a networking event.

Add Tags for Persona

In this scenario, the persona recently started a new marketing role in a medical wearables company. Participants were asked to read their personas and fill out a select 5 interest 'tags' on pre-meetup Google Form that would describe their persona at a Wearables event.

Think-Aloud with Prototype

After a brief training on how to use the HoloLens, the subjects are then taken to a separate room where they encountered three available confederates to talk to. Based on their selection of tags in the previous step, the three confederates were selectively assigned tags that matched to each of the participants' tags. Utilizing a think-aloud process, the subjects expressed their internal dialogue as they choose who they would first talk to and their expectations with the filtering tag UI.

Debrief & Discuss

Afterwards, participants were interviewed on specific actions and emotions, their past experiences in networking events, and how this AR experience might compare with what they previously encountered.

Listening to a participant during the think-aloud portion of the study.

Based on our interviews, we discovered that our participants experienced common issues in current networking events. Participants found it difficult to network with a goal-oriented approach, due to the lack of signalling. Instead, participants networked based on these elements.

Familiarity

A participant once observed that it was difficult for her to meet targeted people at a networking event because they were all in a close group talking with each other. Another participant also mentioned that going to a networking event with friends made it less intimidating, but made it hard for him to meet others, which defeated the point of the event.

Availability

All of our participants have expressed that they choose who to talk to at a networking event based on conversational availability. This is either based on seeing individuals who aren’t currently in conversation, or by seeing an opening in a group. One participant stated that at an event, he chose “tables that were almost full”, and another expressed that she found it intimidating to talk to company recruiters because they were all grouped together in a closed circle.

Homophily

Some of our participants also chose conversational partners based on shared traits. Gender and age were commonly cited reasons, as well as organizational affiliation, such as college alumni.

Objects of Interest

Some participants found it easier to network with others near the food spread and demo booths. This allowed networking participants to interact with objects of interest with others comfortably.

Current Issues at Networking Events

There’s a shyness thing with a meetup. It's awkward, you almost want to go with people to the event, but the flip side is that you just hang out with those people and not break off to meet new people, which is the whole point.

Participant 2

A real-time view of a study participant interacting with the prototype through the HoloLens.

The participant uses the tag-filtering interface to highlight possible conversation partners based on common interests.

Results
from User Study

This is super helpful. I've been to so many events where I only have one hour and I want to meet people I actually want to talk to.

Participant 5

AR helped users prioritize who they want to meet.

All participants expressed positive sentiment from the study. Many participants found it very convenient to immediately find people relevant to their interests. One participant expressed that it allowed her to make a strategic decision about her networking, which replaces her usual method of finding the first person alone and closest to her. Another participant stated that this was especially useful in limited timeframes, because it allowed him to rank who he needed to talk to.

Job titles and affiliation trump matching interests.

We did not initially test for this, but through our think-alouds, all 5 participants prioritized their meetings based on the confederates' job title and affiliation. All 5 wanted to network with the CMO of MedSwift, regardless of how strongly his interest tags matched with their personas.

AR created an opening for conversation.

4 of 5 participants felt that the AR system helped them form their first question for approaching their target conversation partner.

Users were comfortable publicly sharing their interests.

Users expressed no discomfort knowing that their own interest tags were publicly viewable by others. We elicited responses ranging from "I chose the tags, so that's ok", "all these things would be on LinkedIn anyways", to "I think that's awesome" and "I'm pretty introverted to it's helpful to have this".

Study Limitations

Based on our study of a nascent technology, our design had some significant limitations.

Nascent OS

The Hololens is still a new technology with sometimes unpredictable behaviors. We encountered many glitches while developing the system, and had to reset the program for 4 of the 5 participants. Additionally, Cortana once unintentionally activated and would not leave the user alone.

Resource Limitation

Ideally, we wanted to test how multiple HoloLenses interacted with each other. Due to our limited resources, we constrained our study to a single Hololens looking at three confederates.

Limited FOV

The Hololens has an extremely small FOV, making it difficult to be useful in a large, crowded environment. We had to prompt users to look at specific interface elements which might be more immediately discoverable on an interface with wider FOV.

Artificial Environment

User behavior may differ in a lab environment vs. a natural testing environment. In the future, we'd like to test in environments that may induce more realistic feelings of social anxiety or self-consciousness in participants.

Conclusions

Our study showed that AR influenced the conversation partners users prioritized during networking meetups. In past networking events, our participants chose who to talk to based on familiarity, proximity, availability, and superficial homophily (same gender, age, affiliation). Their methods of choosing conversational partners were somewhat arbitrary and irrelevant to their goals at the event.

With our AR system, users were able to prioritize who to talk to based on surfaced information that would otherwise be hidden in a non-AR context. Novice and experienced networkers both appreciated the AR system in reducing their cognitive load; those with extensive networking experience expressed strong positive reactions, noting that this system would have saved hours of preparation and research. Novice networkers expressed relief at a system that would help them remember facts of multiple new contacts. It seems that job position and affiliation outranked matching interests, but matching interests still served as an important function as potential conversation topics.

Interestingly enough, our system also seemed to make users feel more comfortable in other ways. In our interviews, participants discussed the feeling of awkwardness from the ambiguity of being in a room full of strangers. As one subject put it “people feel awkward about starting a conversation when they don't know anyone”. One introverted participant appreciated how his own interest tags might help others approach him.

Next Steps

Our study showed promising preliminary results for bringing AR into a networking event. However, many questions remain unanswered.

Personal Interest Tags

Many of our participants requested more hobby-based tags to create a more "human," less "formal" approach to networking.

3D Placement of Information

Our tests surfaced an interesting issue: where should the AR be placed? Should it stay fixed on others' heads, or relatively positioned to the user? Our tallest participant noted possible discomfort with using our system.

New Entrants

If another individual who is a 'strong match' enters the event after the primary user, how is the primary user notified?

Tag Ambiguity

Participants wondered about the meaning of a tag. Is the matched partner knowledgeable about the topic written on the tag, or curious about the tag as an interest?

Name Display

A participant expressed strong discomfort of potentially having a stranger come up from behind him and call his name with familiarity. Another participant, a former saleswoman, wished she could see the full names of others. Further study needs to be done to test user comfort.

Next Steps
cont’d

Attention

Throughout the whole process, how might AR enhance the user’s goals without impeding them? Our team was not able to test extensively for conversational effectiveness.

Information Density

Although there were many participant requests for ability to see more information, our previous studies show that too much information may actually be detrimental to the conversation, introducing distraction or artificially structuring a conversation.

Queuing

For popular members at a networking event (for example, a guest speaker), how might we structure the AR so that these individuals can better manage the queue of people trying to meet them, and improve the conversational experience of both parties?

Recruit More Participants

Our participants were mainly recruited using convenience sampling from our first-degree networks, which may skew results based on demographics (Masters’ students) and relationships (social desirability bias). In the future, we hope to recruit more subjects in different age ranges, different roles, and outside of our first-degree networks.

The team would like to thank the following individuals for their support of our research:

John Zimmerman
Jessica Hammer
Anind Dey
Brandon Taylor