Attention from Afar: Simulating the Gazes of Remote Participants in Hybrid Meetings

Bin Xu[1, 2] **Jason Ellis**[2] **Thomas Erickson**[2]

[1]Information Science
Cornell University
Ithaca, NY14850 USA
bx55@cornell.edu

[2]IBM T. J. Watson Research Center
Yorktown Heights, NY 10598 USA
{jasone, snowfall}@us.ibm.com

ABSTRACT

Gaze is a powerful form of social feedback, providing cues about attention and interest, and boredom and distraction. We designed a working prototype that enabled remote participants in a collocated meeting to look around the local meeting space, and that showed local participants where the remote participants' "simulated gazes" (that is, their virtual cameras) were directed. Of course, pointing a camera is not the same as gazing, and so we conducted a study to understand how simulated gazes might be used, and to what extent they would be experienced as social cues. Findings range from the use of simulated gaze to signal attention, to ways in which local and remote participants experienced these simulated gazes. These findings illustrate the value of indirection and abstraction in presenting social cues; raise issues of privacy, visibility, and participation asymmetry; and suggest implications for design and further research.

Author Keywords

Hybrid meeting; gaze; computer-mediated communication.

ACM Classification Keywords

H.5.3. Information interfaces and presentation (e.g., HCI): Group and Organizational Interfaces: Synchronous Interaction.

INTRODUCTION

A hybrid meeting is a meeting in which several people are collocated in a meeting room, but that also has remote participants connected by phone or video. Although common, hybrid meetings are challenging for remote participants because of an impedance mismatch: the collocated participants can readily see and converse with one another, whereas remote participants are less visible to collocated participants and their options for signaling their engagement are more limited. In our view, enabling effective hybrid meetings is something of a grand challenge in designing tools for meetings. This is particularly true if one wishes, as we do, to avoid the use of high-end, dedicated video-conferencing rooms, and the need for remote participants to use specialized equipment. In this paper we report on formative design explorations that aim to reduce (but by no means eliminate) the impedance mismatch between collocated and remote participants.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from Permissions@acm.org.
DIS 2017, June 10-14, 2017, Edinburgh, United Kingdom
© 2017 ACM. ISBN 978-1-4503-4922-2/17/06…$15.00
DOI: http://dx.doi.org/10.1145/3064663.3064720

Specifically, we designed a working prototype with two features. It uses an omnidirectional camera in the meeting room to let remote participants independently look around (each participant using their own virtual camera). And it provides abstract representations that show those collocated in the meeting room where the remote participants have pointed their virtual cameras. Remote users only need a device that supports voice and a display to use the system.

In particular, we are interested in whether showing collocated participants where a remote participant is looking can function as a "simulated gaze." Gaze is a powerful social cue [6]. It can signal whether a person is interested or bored, and whether they are following a conversation or not. But of course, pointing a virtual camera is quite different from one human gazing at another, which is why we use the adjective "simulated," and do not assume that the benefits and uses of real gaze apply to simulated gaze. Since neither existing theory nor prior work seemed sufficient to suggest how simulated gaze might work, or how to use it in a design, we developed a working prototype to better understand its usefulness.

This paper describes the working prototype (Figures 1 and 2), the method, and issues and design implications. It begins by reviewing prior work, and then describes the working prototype. Next is the methodology, in which 19 participants used the prototype to carry out a discussion task, and then participated in a semi-structured interview. The results section describes the participants' experiences, their reactions to the system, and discusses our findings. The paper concludes with design implications and directions for further work.

RELATED WORK

Using Gaze in Collaborative Systems

Gaze is a powerful cue during face-to-face conversations [6]. It can silently trigger a conversation by capturing visual attention [4, 22, 26], refer to objects and persons [9, 30], coordinate turn-taking in a group [37], and indicate interest in other persons [43]. Using gaze to communicate one's intentions is a part of human communication [e.g. 24, 25], as is our ability to use it to infer others' intentions [e.g.,1].

Research on collaborative systems has also explored the utility of gaze to support both conversation-oriented and task-oriented collaboration (e.g., [9, 26, 28]). Other work explores ways to allow virtual environment users' avatars to engage in mutual gaze [30], and to allow them to make eye contact in avatar-mediated virtual meetings [5] or an immersive virtual collaboration [19, 28]. Gaze was also studied in video-mediated collaborations. Monk et. al. designed a system to support awareness of gaze in video conferences and found that gaze awareness served as an alternative non-verbal channel for the development of common ground [33].

However, gaze in these systems differs from the real gaze that shapes face-to-face conversation. There, gaze was either captured by gaze-tracking devices (e.g., [9, 48]), or delivered through a 'transparent' interface that aligned users' viewing directions (e.g., [26, 28]). There are pros and cons to these methods. Although gaze-tracking devices are accurate, they must be calibrated before use and aren't ideal for daily use [44]. Solutions like [28] are less complicated, but are still limited to specially configured physical spaces.

Our aim is to explore approaches that are cheaper and less technically complex, and therefore have the potential to be broadly used. Our starting point was to deploy pan-tilt-zoom cameras [41] in a meeting room to empower remote users to look around, but these cameras provided a single shared view to all remote participants. As the number of users increases, control of such a camera is still limited to a single person and conflicts arise if there is disagreement about the direction it should be pointed. This limitation inspired us to use an omnidirectional camera [39], which uses wide-angle lenses to capture a sphere and, in this way, can "see in all directions at once". The camera streams this panoramic video and enables each remote participant to move a virtual viewport to a part of the panoramic image. This led us to wonder whether these individual virtual camera movements could stand in for the gazes of individuals, though the simulated gazes are not as accurate as what can be captured by eye-tracking devices.

Once gaze, real or simulated, is detected, we must consider how that gaze might be presented to others. One solution is to represent the gaze by animating the eyes of a users' avatar [5, 19]; another is to use a physical device whose orientation corresponds to the user's gaze, for example, by using a small rotating screen that depicts the remote person's face [31, 32], or a telepresence robot.

Our aim is to find an affordable way to present remote users' simulated gazes to local participants. While using avatar eyes and head movements to depict gaze has the advantage of mimicking natural behavior, we are not yet at a point where either virtual or augmented reality systems are being commonly used. Telepresence robots are somewhat more common, but are expensive and far from ubiquitous. Furthermore, availability aside, although telepresence robots can depict gaze, the use of telepresence robots does not scale to situations where there are many remote participants – physical meeting rooms can quickly become crowded, especially because maneuvering telepresence robots is still, at best, cumbersome. Inspired by [42], which shows how an abstract visual presentation can represent user status and activity, and following a discussion in [7], which cautions that a full and detailed presentation of remote person could be privacy sensitive and distracting, we decided to explore representing gaze with an abstract but hopefully legible visual representation.

Social Awareness in Collaboration

Awareness has been an important research theme in Computer Supported Collaborative Work (CSCW) since its early days [10]. Awareness in shared physical and virtual workspaces includes knowing who is present [e.g. 11], and other social factors [15], including knowing what others know [38], who they are connected to or their social network [53], and what they're doing (e.g., [47]). Research also shows the benefits that these kinds of awareness can provide, improving the performance of work [52], the coordination of tasks [29], the creation connections between employees [40], and team dynamics [8, 29].

Awareness is also important for distance collaboration. Early work like "Portholes" demonstrated the positive effects of having awareness of others when collaborating over distance [11]. Specifically, increasing a person's awareness of who the remote collaborators are provides an opportunity to enhance the presence of these remote users, even without embodied presence [42]. Approaches to supporting awareness across distance include sharing the video stream of a remote space and persons [11], sharing the planning and coordination activities of remote teams [40], and information about remote user activities [7], etc.

In hybrid meetings, local persons share a meeting space and are aware of each other's presence and attention, but remote participants have less access to this information. We expect that showing the simulated gazes of remote participants may be beneficial for awareness.

Social Translucence and Privacy

Making activities transparent has benefits and risks. The social translucence framework provides ways of thinking about how socially salient information can facilitate or inhibit social processes in collaborative systems [12, 13] and social media designs [21]. In a system that supports independent views of hybrid meetings, one question is whether all collaborators need to see the same things to build common ground. Limited work discusses this question in terms of gaze sharing. In this paper, we examine issues of social translucence by adopting a similar design.

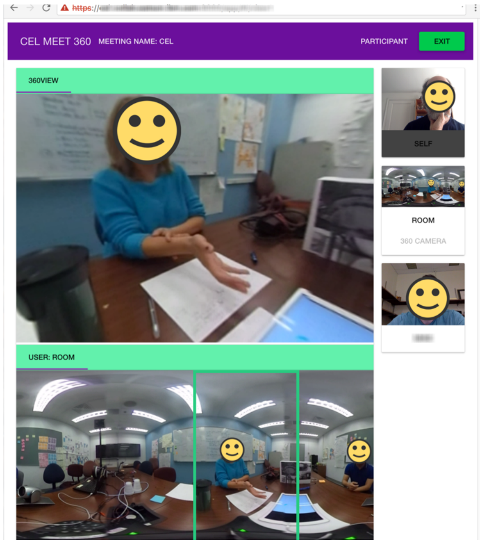

Figure 1. The remote user's interface: The upper panning view allows users to look around the meeting room; the lower view shows a panoramic video of the entire room – and the green rectangles show what the panning window is viewing. The right margin contains thumbnails of the views from all sites – the meeting room (middle) and two remote users (above and below).

Because gaze indicates attention and interests [30], it may raise privacy concerns. For example, one study showed that employees felt their privacy was violated because they, their activities, or their data received unexpected attention [3]. Therefore, in hybrid meetings, it's possible that inferences based on the simulated gazes could lead to privacy problems, social pressure, or other forms of discomfort. We hope to gain insight into these issues.

PROTOTYPE SYSTEM

To support hybrid meetings, we developed a working prototype with two components: a Remote User Interface to allow remote users to view the meeting room, and a Gaze Simulation Interface to display remote 'gaze' to collocated users in the room. We also built a logging system to produce Usage Charts used to facilitate discussion during the study. These components are described in turn below.

Remote User Interface

To allow remote users to view the meeting room, we used an omnidirectional camera [39] to capture the space. The Ricoh Theta S camera [46] was chosen due to its low cost, popularity, and live streaming capabilities. The camera produces a 1280 x 720 live stream at 15 fps. We captured this video and streamed it using the WebRTC protocol.

The Remote User Interface (Figure 1) has three components: panning view, panoramic view, and remote participant thumbnails. The panoramic view (bottom) shows the entire image captured by the camera as a rectangle. This panoramic image is distorted at the top and bottom. We positioned the camera to make the faces of seated people fall along the equator of the image, so faces

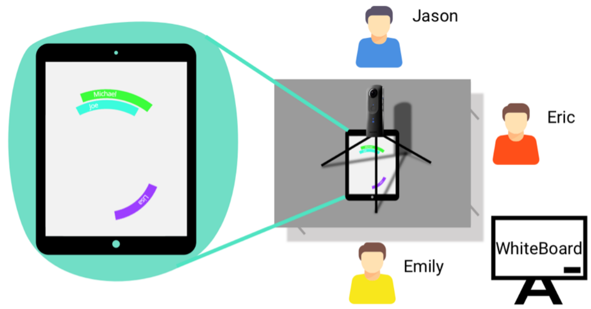

Figure 2. Gaze simulation as seen by local users. The gaze simulation interface sits on the table, adjacent to and aligned with the omnidirectional camera. Each arc shows where each remote person is looking using the panning view. Here, 2 users are looking at Jason and 1 at the whiteboard.

are typically recognizable, while higher (ceiling) or lower (conference table) objects tend to be distorted.

To compensate for this distortion, we provided a panning view (top in Figure 1) that shows an undistorted portion of the total video sphere. Users use their mice to control the direction of their viewports (aka virtual cameras). This functionality allows multiple users to look around the room in real time, independently and simultaneously. We call this ability to control the virtual cameras "simulated gazing" and use it to drive the Gaze Simulation (described below).

Lastly, the Remote User Interface captures video of the participant using her/her laptop camera and streams to the other remote users. Each of these video feeds is shown in a thumbnail along the right side of the user interface. When a user clicks on one of these thumbnails, the video expands to fill the large video window at the top left of the interface.

Gaze Simulation Interface

With the remote interface, the system captures where remote users point their virtual cameras. We also needed a way to display these simulated gazes in the meeting room. As discussed, we used an abstract representation to achieve this (Figure 2). This representation – or gaze interface – displays an arc for each remote user, the arc corresponding to the horizontal direction in which the virtual camera is pointed. As remote users change where they are looking, their arcs move accordingly. The simulated gazes were displayed on a 12.9" iPad on the table next to the camera. Local participants could view the gaze interface and see where remote users were looking by watching the arcs. Figure 2 illustrates a case where two remote users are looking at Jason and a third at the whiteboard. Note that the 'gazes' presented here are not perfect. To simplify the interface, only horizontal movement is represented. Thus, if a remote user looks up or down, this is not reflected.

Logging System and Usage Charts

For the purposes of the study, we also built a system to log each person's interaction with the panning window, including each time they changed their viewpoint and how

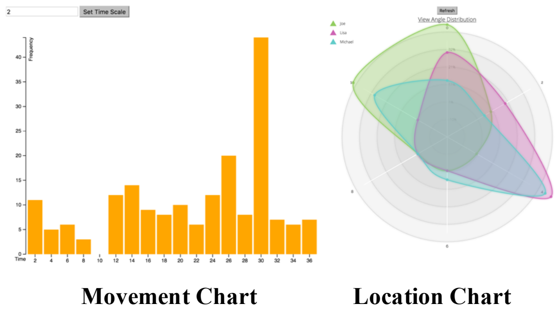

Figure 3. Usage charts. Movement chart shows how many times the remote user changed their perspective in the panning view. The location chart shows how long the user maintained a particular point of view in the panning view.

long they stayed at each viewpoint. These data are used to generate two usage charts – the movement chart and the historical location chart (Figure 3) – that were used in the study. The movement chart shows when and how many movements a user made during the study session, and the location chart shows the frequency with which a user looked in various directions. The logging mechanism was a Google Chrome web application, the client was built with HTML5, AnguarJS and D3, and server was Node.

METHDOLOGY

The Study

The purpose of this study was to determine how users used and experienced the various interface components, and to gather ideas for future design. In brief, 4 people (2 collocated and 2 remote) carried out a discussion designed to permit them to exercise the features of the technology. Afterwards a semi-structured interview, augmented by usage charts, was used to gather feedback. Details follow.

Setup

Because the aim was to understand both the experiences of those 'in the room' and those who were remote, each study session included two 'real' participants, P1, the local participant and P2, a remote participant (i.e. in a nearby room but connected only via the system). Each session also included two Experimenters: E1, who was in the room with P1, and E2, who was remote (but in a different location than P2). As noted in the description of the system, remote users E2 and P2 could independently control the direction of their views in the panning window of their user interface. A schematic of this setup is provided in Figure 4.

Task

Participants were told that they were members of a small group that had been asked to decide how to spend $100,000 to improve the social/work environment of the building in which they worked, for one of three groups (visitors, remote employees or regular employees). We designed this task to provide reasons for participants to both look at one another, and to look at objects in the room (e.g., a map). In addition to the design of the task itself, the 2 experimenters also acted in ways to encourage P1 and P2 to exercise

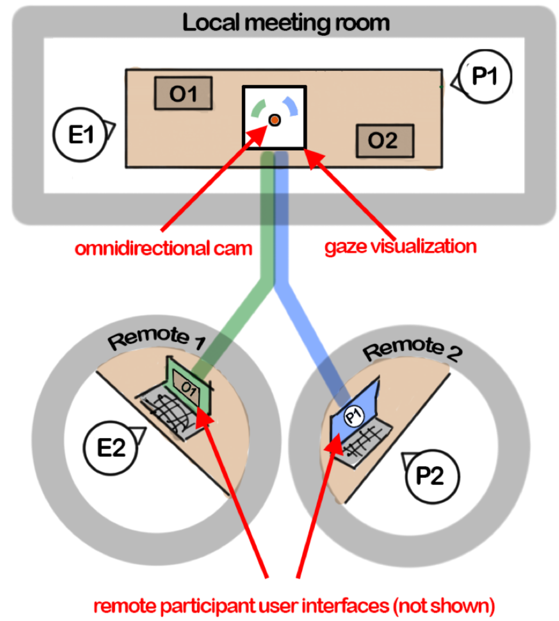

Figure 4. A schematic of the study setup: An experimenter (E1) and participant (P1) collocated in a room with an omnidirectional camera (center) and various objects (O1 and O2) on the table. A second experimenter (E2) is in one remote location, and a second participant (P2) is in a second remote location. E2 and P2 can independently adjust their views to look around the meeting room.

features of the interface. The first part of the discussion was guided by the two experimenters so that participants talked to (and optionally, looked at) the others. In this part, E1 acted as the coordinator of the discussion (and revealed himself as an experimenter), while E2 acted as a confederate who was a peer of P1 and P2. Typically, this discussion focused on which of the three groups should be supported. The second part of the discussion was guided to focus on changes that might improve the work environment. Here the E1 and E2 began referring to objects in the room, such as a large poster that showed areas of the site that might be suitable for alteration. After the first two sessions, other objects were added – e.g., a telepresence robot – to allow for a richer discussion of improvements.

Timeline

Participants P1 and P2 arrived at the study location and were introduced to one another by E1. The purpose of the study was described, the technology was demonstrated to both, and consent forms were signed. Next P2 was taken to a separate room and set up with the remote users' interface (Figure 1), reminded of how to use it and that the people in the room could see the direction in which the panning window was pointed. Third, E1 took the other participant, P1, into the meeting room. At this point, the discussion task began: it lasted 15-20 minutes, and followed the two-part discussion format already described. Fourth, the discussion was ended, and P2, the 'remote' participant, was brought back to the meeting room, and a group semi-structured interview was carried out with P1 and conducted by E1.

Semi-structured Interview

Each session concluded with a semi-structured interview. Initially, participants were simply asked to report on their reactions to the system; after about 10 minutes, they were shown the two Usage Charts (Figure 3). These charts were used as probes: first to extract more detailed accounts of participants' actions/reactions, and second as prompts to discuss privacy concerns.

In general, the interview covered the following topics:

- Meeting support technologies participants has used
- The remote user's experience of the remote UI (Figure 1)
- The local user's experience of the gaze simulation interface (Figure 2)
- Discussion with both users about when and why they made use of (or did not use) their portions of the UI
- Discussion with both users about the usage chart (Figure 3) and whether it should be private or shared
- A discussion about other uses for this sort of technology

Participants

Since the aim of the study was to gather reactions to the technology, and collect ideas for further development of the system, we sought technologically adept participants who had prior experience that could inform their comments. No attempt was made to recruit "average" participants to assess issues like ease of learning, usability, or likelihood of adoption. In our view, the prototype system was in too early a stage for this to be of value. Therefore, this study recruited employees from the Research division of the company in which the technology development was occurring. The company itself is a global information technology firm whose businesses range from hardware development to services provision, and which has on the order of a hundred thousand employees. Participants were recruited by approaching people in the cafeteria, who were later contacted by email to confirm availability; snowball sampling was used to expand the list of candidates.

We scheduled 10 sessions with 2 volunteers each. 1 volunteer did not show up, so in that session we only have the remote confederate as a remote participant. There were 19 participants (15 males and 4 females aged from 20 to 60+). All took part in the final interviews.

Data Analysis

Interviews were audio-recorded, transcribed and edited to remove identifiers and other references that may identify the participants and/or anyone they mentioned during the interview. Quotes from participants who served as "local users" are numbered (L#), and quotes from those who served as "remote users" are numbered as (R#).

After reading the transcripts to become acquainted with the data, transcripts were imported into the Dedoose qualitative data analysis tool and divided into meaningful units. Three of the authors did a close reading of the transcripts while identifying key points, as a part of an open-coding process in which we coded distinct concepts and categories in the data. After this, they examined relationships between key themes to refine categories, and to ensure close associations between participants' responses and emerging analyses.

RESULTS

Four categories emerged from participants' feedback: the usefulness of two different views in the remote interface; the meaning of the simulated gazes; the effects of asymmetry in interfaces and data; and the privacy concerns around collected data in system. We discuss each in turn.

Panning and Panoramic Windows Have Different Uses

Participants in the remote role saw two video views: the panning view that they could rotate to look around, and the panoramic view that showed a distorted but omnidirectional view of the conference room.

Panning view offers immersive perspective. First, the panning view was consistent with users' visual perception of space, as it "*is a pretty standard view that you get in any conference system* [and] *what I expected.* (R10)", and offers better *"spatial sense throughout the discussion* [that] *is very helpful*. (R7)*"* Especially when *"sometimes* [participants] *talked about some things in the room, the objects, and* [remote participant] *can definitely move around and focus and look at that object immediately.* (R7)*"* And it further increases remote's sense of belonging: "*It mostly helped me feel more included. It's not like we're talking and I have really no idea what's going on in the room.* (R6)"

Remote participants also found the panning view useful when there was significant turn-taking in the conversation in the local room, or those in the meeting room interacted with local objects: "*I changed the direction when there was a transition between speakers. And also when you point to a specific object in the room on the table then I tried to look at that object. I think that was an advantage*. (R7)"

Panoramic view offers navigating reference and overview. The panoramic window was used to "*follow what was happening there as people were pointing to, then you would want to look at that.* (R4)", and "*see overall who and what was there and sometimes I looked at that instead of moving the top camera to the person.* (R3)"

Regarding offering two views in one interface, participants reported it was beneficial to have reference/overview alongside a more detailed panning view: "*I think the combination of the views and helping with that, because ... When you have the* [panoramic view]*, you see everything. If you just had the portal view* [panning view] *it's like, I'm looking at him but at the same time you're doing something which I'm totally unaware of because I don't see it.* (R6)"

However, other participants tended to prefer just one view. When there were fewer turns in the conversation, the panning window was less used, and they reported looking at the panoramic window briefly "*was sufficient to give an idea of what was, where you guys were, and what things were around and stuff* (R4)" because: "*all that was* *happening was people were talking. I could focus in on you* [in the panning view]*, rather than on the bottom I could see you both, right? That doesn't really give you that much additional information.* (R4)" Some other participants preferred to have just the panning view (top) because it was not distorted and "*was very responsive and quick to just have a look around with it, and the lower* [panoramic] *one gives you that overview but it was a bit distorted.* (R3)"

Overall, the usefulness of being able to look around was highly related to whether there were *contents* located in the space, for example, an object of interest (e.g., pointed at by a local person), or *contexts* important to the conversation itself, for example, non-verbal cues of local people [49].

One way to understand the results is to think about how and when the two views provide information relevant to the conversation. The hyper-personal communication model [50] suggests that in digitally mediated communication, people are still able to achieve common ground without the non-verbal cues that face-to-face offers. This may explain why some remote participants expressed less interest in moving their viewpoint to the talking local person when the discussion was primarily verbal. When participants began referring to objects in the room, seeing the video became more important to following the conversation.

In summary, the panning view offers a non-distorted view of local room, which helped remote participants make sense of the space and focus on interesting activities and objects, which is hard to achieve through verbal conversations alone. The panoramic view is better for getting an overview and providing reference for the panning view, especially when there is a need to move the panning view. When participants were in a primarily verbal conversation, the panoramic window was quite effective on its own.

Simulated Gaze: An Indirect and Abstract Attention Representation Has Advantages and Disadvantages

During the discussion, participants paid occasional attention to the gaze UI, but did not concentrate on it as it "*was just* [in my] *peripheral vision* (L10)".

The movement of arcs in the gaze UI however did attract attention. Seeing an arc move led to "*inferences that maybe you* [remote person] *see this, and maybe you're following it, and that you made some attempt to move your camera around to look. That indicates some level of attention that you're doing something to look over in that direction* (L4)". Beside movement, "*if there is more than one person and they are looking in the same direction, then the people here can recognize that there is attention coming from the remote place as well.* (L7)"

Awareness of remote attention offered benefits to local participants. Besides being "*entertaining to watch* (L6)", "*it's not intrusive and it's easy to have in a situation. It doesn't necessarily always have a use, but when it doesn't have a use, it's not that annoying.*(L4)"

The simulated gazes also lead to the local person having a heightened sense of being visible and being paid attention to, or not: "*I think initially I was a little bit more exposed. People are turning, looking, and you're like, 'Oh, boy, better put on my best for the teleconference,' you know, I'll be eating granola bars and whatever else while the conference would be on, ... when the direction of the camera, of the focus, was being pointed away even though you've got that other overlay image at the bottom but it was pointed away, I sort of relaxed a little bit more.* (L6)"

Having a gaze simulation UI in conference room also increased the remote person's self-consciousness about their performance. They tended to adopt the norm that they should "*pay attention to be polite* (R2)", as well as to demonstrate their commitment to a meeting: "*if I was in a meeting that was an important meeting or a meeting with a lot of people, it's a way to show that I'm participating. That I'm awake and listening and following the conversation and not just off doing something else* (R3)". And such social pressures were also seen as a way to increase engagement: "*This is nice to give some feedback in terms of whether* [the 2 remote users] *were paying attention to what we were saying. They wouldn't be able to get away with pressing the button on mute at home and dancing around in their pajamas because if it's not moving and looking at where we're pointing to say, 'Hey, let's talk about this topic,' you knew that they're not paying attention. It puts pressure on everyone to be fully engaged* (L6)".

From a Goffmanian perspective [22, 23], making a remote person's attention visible brings the remote person more onto the front stage than a traditional video window that provides no gaze information (i.e., Mona Lisa Effect [34]). Being on the front stage stimulates the self-consciousness of performance and self-presentation, and increases social pressure to behave in particular ways. This also may explain why people in the conference room, who are already on the front stage, now get a better sense of attention from the remote audience and, thus, behave as if on stage. Without presenting simulated gazes, local persons may assume that remote audiences are always or never paying attention, which is a similar effect to that observed with self-presentation on the Facebook Newsfeed [2, 14].

However, the simulated gaze is not fully equivalent to real gaze, and local participants were aware of that: "*they* [remote persons] *might be looking at the lower window. Where they can see people at once and then they won't move* [the panning window to trigger arc movement]. (L3)" The ambiguity regarding whether the remote person was paying attention or not led to positive attributions when users were not following the norm of looking at people to be polite or show engagement: "*I could use this* [simulated gaze interface] *as an indicator of where people are looking and I saw* [remote person 1] *being much more active than* [remote person 3]. *I mean there's just something in the back of my mind wondering you know what's going on with* [remote person 3] *why aren't you looking around and the answer is you were just using the panoramic view.* (L9)*"*

Taking the perspective of the casual attribution theory [51], our prototype, as a socio-technical system, injects the ambiguity that simulated gazes do not always reflect what others' are attending to. This is similar to the finding in a machine translation mediated collaboration [20]: when there are two possible interpretations of communication errors, one due to system limitations and the other one to the human, communicators attribute errors to the system.

Finally, remote participants found that seeing simulated gazes helped improve their feeling of presence in the conference room: "*When I'm in the room, I can see, 'Oh, you know* [remote person] *is looking around,' and you can come in the room and say, 'Oh, hi* [remote person]*.' It's kind of nice. It's kind of sitting there. There's obvious value for remote people. … Then also, for people coming in the room, they're kind of aware there might be these other people who are not there in person but who care about what's going on.* (R6)"

In short, the interface presenting simulated gazes is seen as an abstract view that provides a partial representation of the remote person's attention, and it has both advantages and disadvantages. The advantages are that it is less intrusive, it provides the local person with awareness of the remote person's attention, and it increases social presence by signaling the remote person's participation and activity. At the same time, it can increase social pressure on the remote person to pay attention, and may increase the local person's concern about receiving too much (or not enough) attention. Finally, the fact that the indication of attention is not entirely accurate, can lead to misinterpretation, but can also provide ambiguity that can support plausible deniability.

The Viewing Asymmetry Has Pros and Cons

As discussed, we were inspired by the social translucence framework to create an abstract representation of gaze. We chose not to show video of remote participants in order to let participants focus on the gaze interface rather than overwhelming them with multiple video streams. However, remote users were able to see video of the local room and other remote users. In contrast to the asymmetry in traditional meeting setting where local persons have more access than remote ones, our prototype introduced an asymmetry that limited local persons' access: "*I felt jealous of* [remote] *people who had cameras, who had a display. I could not see* [remote #1] *or* [remote #2] *and I felt a sort of asymmetry there. I felt blind where they could see so there was a kind of jealousy there of their ability to see more than I could* (L9)", or as a sense of inequality: "*People need to feel equal… and that's the best way I can describe it, otherwise it feels creepy.* (L7)" One solution the participants suggested was to integrate each remote person's video into the gaze interface: "*I really want to see the person, not just have their name. I mean the ideal thing would be if that was the image at the top*. (L10)" Although we deliberately designed the gaze simulation interface without the live video to probe reactions to simulated gazes, it would be useful to explore ways of combining gaze with video, etc., as we move towards a more functional design.

Despite the asymmetry, local participants experienced benefits from this design choice. First, by presenting remote's viewing behavior as an abstract visual cue to participants in the local room, the prototype required low cognitive load as the design *"makes it easier to understand out of the corner of my eye* [and] *this conveys less information and it requires way less cognitive bandwidth from me.* (L9)" Another potential benefit is that the gaze simulation is low fidelity and so avoids potential privacy issues or distraction often associated with live video: "*It's also good in some sense it's a low fidelity interface. It's more lively than, 'here are a list of people who are logged into the camera.' … it's giving some sense of where people are looking, but it's kind of low fidelity. I don't think it's revealing a lot of stuff about you and I think it's a good thing to have.* (R6)"

Making Sense of Privacy, Making Sense of Data

Another asymmetry that the prototype introduced was in what data was collected. That is, remote participants' movements of the panning window are captured by the system, and potentially available for further use, unlike the collocated participants' gazes. Two usage charts (Figure 3) were used to demonstrate potential uses of the gaze data. They summarized the remote person's movements over time during the (movement chart) and the degree to which they looked in specific directions (location chart). The main purpose of visualizing these data was to let users to reflect what had happened in the meeting and help us examine to what extent this kind of data can help users understand the meeting and their own performance.

Remote participants felt that the two usage charts provided interesting insights on their own performance during the meeting: "*if my pattern was anomalous for me or different from most people and the system* [charts] *were to tell me that it might be an interesting: 'Are you OK, were you especially excited or bored by this meeting?' If it were an entree to a further exploration of why were you behaving in this way, it might be interesting.* (R9)" and there was value to "*understand my behavior a little bit better.* (L3)"

However, participants varied in their opinions about whether the charts should be shared. Some felt it was OK to share as the charts were just summaries of publicly observable activity, which "*is not really private because everybody saw it at the time, if they had been watching* (R3)". Others disagreed. The first concern is that there was insufficient context to allow others make sense of the gaze data: "*I don't know where they've moved to so there doesn't seem to be enough information there to care about this. It's not correlated with activity or anything* (L10)". Further, it also missed the important contexts like "*what was being said and who was saying it, what was being referred to and* *stuff like that*. (R7)" And the conversation itself "*is a context that is not being captured that is important to interpreting and giving you valuable insight.* (L4)" This in turn could result in misinterpretation and privacy concerns: "*If I interact with a person and that person never looked at me then I would say if that person really care about what we are talking about and so forth. It does have some implications, privacy implications*. (L3)"

Another concern was that sharing the usage chart would make remote persons self-conscious: "I am *going have to be self-conscious about where I look all the time and is this going to make it really obvious who talks and who didn't talk in the meeting and am I going to be the one person in the meeting that nobody points at?* (R3)" The same participant also suggested sharing the historical data might lead to behavior changes: "*I think I would probably behave differently if it was or wasn't shared…. If it's not shared the whole thing is more like a telephone call where you have a certain anonymity and you can do what you feel like and pay attention to what you. ... Where if it's shared people are going to draw conclusions from it.* (R3)"

One way to understand the importance of contexts is through the lens of Nissenbaum's context integrity framework [35, 36]. The attention, whether inferred from the simulated gazes or observed by others, co-occurs with a context of an episode of the meeting: the topic, the leader, and what has been discussed. The quantification of remote persons' simulated gazes, regardless its accuracy, loses the contextual and conversational information that supports interpretation. Later inspections of this data might use a default context – e.g., what a 'normal' meeting looks like – to interpret it. The potential for misinterpretation raises concerns on the part of data owners (remote persons). This is similar to what happens in Facebook: its sparse contexts, resulting from interface and access control features, tend to trigger violations of content sharing norms [27].

Participants had fewer privacy concerns regarding showing of simulated gazes in real time. One reason was that the gaze interface is not presenting historical data and *"because people aren't sitting there staring at it the whole time. ... I can just move it at some point and then when anybody looks at it they will see oh it did move since the last time I looked so she's still there.* (R3)*"*. The fundamental difference between the gaze interface and the two usage charts is that the simulated gazes are presented ephemerally: it only shows the present data in the current context, while the usage charts show historical data without context. The embedding of simulated gazes in the current context is similar to the way Snapchat's ephemeral communication design impacts the users' self-presentation [54]. Thus the real-time presentation of simulated gazes is more suited to the communication goals in the current context: *"As a participant in a discussion, what I'd really like is for the discussion to proceed as naturally as possible. For people that means knowing who is talking, who is paying attention to you, who is off doing their email or who is point where. In retrospect* [the usage chart] *is interesting to me but it is kind of a distraction and it is not going to help me towards the real focus of the meeting which is to address the goal of the meeting.* (L7)*"*

Figure 5. A multi-foci + context view of the local meeting environment.

Summary and Limitations
The aim of our study was to understand how the prototype, in general, and simulated gaze, in particular, will be used and experienced in hybrid meetings. Our findings were as follows: The ability to view remote spaces is more useful for seeing the space (and objects in it) than for tracking conversation. Simulated gaze is useful for signaling that remote participants are present and paying attention – all participants recognized this as helpful for increasing social awareness. Simulated gaze created some anxiety among producers and recipients of 'gaze,' – this is another indication that simulated gaze is a powerful cue, even as it indicates the need for more work. Users also agreed that their aggregated gaze data was useful for reflection and research, but were concerned about misinterpretation if it is more widely shared. Finally, we found that visibility and asymmetry of participation has both pros and cons, suggesting the need for more design work.

Our study had limitations. Because it was the first-time participants had used such a system, there might be novelty effects and, as participants mentioned, their understanding of simulated gaze, behavior, and norms might change over time. Also, the prototype was designed to support a hybrid meeting with a single collocated site (a hub), and distributed participants (spokes). Another limitation lies with methodology: although we collected some log data of their use of the system, and recorded participants' conversation as observational data for later content analysis, we chose to focus on qualitative analysis of interview data due to the limited participant sample size. A future study of a more mature system might involve more subjects in order to allow for quantitative analysis. These limitations suggest further research such as a long-term field study of a prototype with a more advanced design, with more data for both qualitative and quantitative investigation.

DESIGN IMPLICATIONS

The Panning and Panorama Windows
Most of our participants valued both windows, though when forced to choose one or the other their choices differed. What is significant is the reasons for their choices. They valued the panning window because it enabled them to focus on people and particularly on objects. Participants

were dissatisfied with the panning window because it did not provide much better resolution than the full panorama window, and several commented on wanting to be able to zoom in, especially to read things written on the whiteboard during meetings. Participants valued the panorama window because it gave them a sense of the whole space (even though parts were quite distorted) and it served as a frame of reference for the panning window so they could see what part of the room they were viewing.

This suggests that a more mature design for supporting remote viewing should provide the following abilities:

- higher resolution images of people and objects
- the ability to zoom in (and out)
- the ability to see the larger context of the space to provide a frame of reference for the close-in views.

Higher resolution can be achieved by using a better camera system, and that would support zooming as well. The ability to see the larger context for the panning window is provided by the current two-window interface, but it also seems prudent to explore a more compact solution using the focus+context pattern as in [16] and [45].

Some participants also commented that in a very active discussion, where conversational turns are rapid, moving back and forth between participants could be cumbersome. This suggests another design implication:

- provide the ability for remote participants to follow rapid exchanges between local participants.

One approach to this would be to provide an automatic mode where the camera could identify who is speaking and shift the view point between people. This, of course, would introduce another form of 'indirection' to 'gaze', in that sometimes the remote participants would be shifting their 'gazes' and sometimes it would be automatically managed by the system. We favor a manual multi-focus + context solution, where remote users would be able to establish multiple foci (a sketch of this concept is shown in Figure 5).

Lastly, our participants saw the value of high resolution (e.g., for reading the whiteboard) but were also aware of the potential for privacy violations (e.g., zooming in to read someone's private notes). One solution would be to explore ways of automatically limiting high resolution viewing to certain zones; another would be give collocated participants a better sense of what remote participants were seeing, thus making remote participants more accountable.

Gaze Simulation and Presentation

As our findings suggest, simulated gaze is a useful and powerful social cue, even though it differs from real gaze. For the local participants, the simulated gaze provided useful indications of the presence and activity of the remote participants. Local participants were generally quite aware of the indirection involved – they understood that they were seeing the direction the camera was pointed, and that that did not necessarily mean that they were being gazed at by the remote participant. Local participants realized, for example, that remote participants might forget to shift their view, or that they might simply have stepped away from their computer. Some also realized that they could not detect the remote person's angle of view – whether they were looking straight at the participant, or down at something on the table, or even up towards the ceiling. This led to some discomfort – one participant seemed to find the ambiguity about whether they were being gazed at or not more uncomfortable than the certain knowledge they were the subject of gaze. This recalls Goffman's concept of "civil inattention," [23] in which strangers in close proximity demonstrate that they are aware of one another (e.g., exchange brief glances) without either staring or completely ignoring one another. While Goffman's concept was developed to explain how a person might have a sense of privacy even in the midst of a crowd, it would be interesting to further explore how it might be adapted to less anonymous events like meetings.

What local participants found most salient was when the simulated gaze moved. This was both because the movement attracted their attention and because it was then that participants were confident that the remote person was paying attention – especially if the movement made sense (e.g., moved to an object being discussed). Local users also wanted a better understanding of what the user gazing at them was doing, particularly when the gaze moved.

As noted in the findings, remote participants also had feelings about the visibility of their gazes. In addition to feeling more present and more involved, they also reported feeling more self-conscious. Some commented that they felt more accountable for their behavior, and that looking at the person who is speaking is the polite thing to do. Several had the experience of forgetting to shift their view and then feeling embarrassed for seeming to 'stare' at another person (another instance of a violation of "civil inattention"). Remote participants also weren't sure what to do when not using the panning window: the gaze indicator, taken literally, would show them staring fixedly at something.

These findings suggest other implications for the design of future versions of the gaze interface:

- distinguish between 'gazes' that have just moved, and those that have been still for a while
- provide a means to detect whether remote users are in front of or looking at their displays
- create a neutral point for the gaze interface, so that it is evident when the panning window isn't being used

One approach would be to enrich the simulate gaze interface. For example, a simulated gaze might brighten when moved, and then gradually fade out until moved again. If face recognition or gaze tracking were implemented on the remote participants' devices, their simulated gazes could be made more salient when they were present and watching. Additionally, or alternatively, if users did not move their gazes for a while, their views fade out, along with the corresponding simulated gazes.

Asymmetry of Functionality and Data Interface

The system we designed was deliberately constructed to provide asymmetric awareness – to allow remote users to see those in the room, but not to provide video in the other direction. This was because we wanted to isolate the impact of the abstract representation of simulated gaze, without conflating it with video feedback. However, as noted, some local participants were uncomfortable about being the subject of the gazes of people they could not see. This suggests one other implication:

- Provide a way for local participants to get a better sense of who is watching and what they are doing.

This would also address the desire of local participants to understand what the remote user was doing when he or she shifted her gaze to them.

This might be done by providing a small composite video of the audience, or perhaps by adding a transient video of the person who most recently shifted their 'gaze'. The sketch in Figure 6 depicts an enhanced gaze simulation UI where rays show the direction of five remote gazes, their length shows how recently each gaze direction was shifted (rays initially touch the periphery and then shrink), and an image of the user who most recently shifted their gaze is briefly shown. We suspect that this last feature would need careful adjustment to strike a balance between informing participants without being disruptive.

While simulated gaze is useful to participants, it also offers the potential for tracking the fine-grained behavior of remote users. While, participants recognized the value of usage data for self-reflection and for research, they were concerned about privacy. In addition, they noted that the usage charts did not have sufficient context to enable full interpretation, and that other information about participants, interaction, meeting topics, etc., would be required for accurate interpretation. This suggests one more implication:

- Provide remote users with a private interface showing their personal gaze data with contextual information.

One approach would be to store users' data into a personal account that is private by default. The data would be saved with other information like the meeting topics (maybe from their calendar application), participants, and where the participants sat in the room (e.g. screen shot of the panoramic view in the remote interface as a reference for the direction of their accumulated gazes).

FUTURE WORK

We used an omnidirectional camera and system prototype to allow remote users to look around a meeting space, displaying their simulated gazes to those in the room. The accuracy of this approach to gaze differs from that provided by other methods (e.g., eye-tracking). It is not clear how much accuracy matters for providing social cues, without explicit comparison to other implementations, but the results suggest that it may have effects on interpretations of what the 'gaze' means. Thus, one area for future work is to

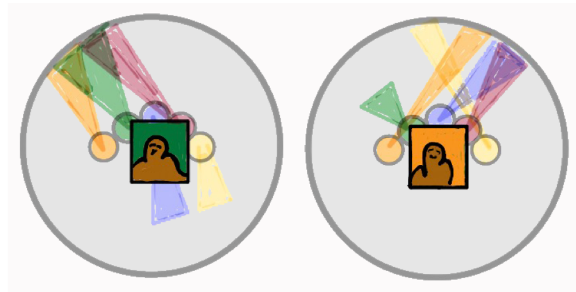

Figure 6. An enhanced gaze visualization. Rays show the direction of gazes, and their lengths reflect how recently each was changed (initially they reach the periphery, and then shrink). An image of the person who has most recently shifted their gaze is briefly shown.

extend our setup with more powerful cameras, and to employ eye-tracking and other methods to understand how other approaches to gaze capture impact its interpretation.

Participants also mentioned that the gaze interface was limited to showing only horizontal gaze. In the future, we are interested in exploring how we might provide a sense of vertical gaze in order to remove such ambiguity. However, care must be taken to avoid overcomplicating the presentation of simulated gaze and increasing the cognitive load required to make sense of it.

We are also interested how to attentions between remote participants, which could be a way to raise consensus between remote people.

CONCLUDING REMARKS

The work in this paper makes three contributions. First, it demonstrates a low-cost working prototype for allowing remote users to participate more fully in hybrid meetings, using virtual cameras to direct their 'gazes' at people and objects in the meeting. Second, using the prototype as a probe, the paper demonstrates that presenting "simulated gazes" can provide powerful social cues, in ways similar to that of real gaze. As with real gaze, our results show that simulated gaze can provide valuable information, or cause discomfort (or sometimes do both at the same time). Third, we draw upon our results to propose design implications, for improving the functionality of the prototype, and particularly for making the gaze simulation more helpful and less discomforting. Having verified the usefulness of simulated gaze, our future work will focus on creating a more robust and functional prototype for field deployment.

We hope that our work will provide a basis for others to pursue this line of research, as we see applications for this functionality not just to hybrid meetings but to areas like education (e.g., students remotely participating in a class) entertainment (e.g., a remote audience for a performance), and other domains.

REFERENCES

1. Geoffrey W. Beattie. 1981. A further investigation of the cognitive interference hypothesis of gaze patterns

during conversation. *British Journal of Social Psychology* 20.4: 243-248.

2. Michael S. Bernstein, Eytan Bakshy, Moira Burke, and Brian Karrer. 2013. Quantifying the invisible audience in social networks. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '13). ACM, New York, NY, USA, 21-30. http://dx.doi.org/10.1145/2470654.2470658.
3. Jeremy P. Birnholtz, Carl Gutwin, and Kirstie Hawkey. 2007. Privacy in the open: how attention mediates awareness and privacy in open-plan offices. In *Proceedings of the 2007 international ACM conference on Supporting group work* (GROUP '07). ACM, New York, NY, USA, 51-60. http://dx.doi.org/10.1145/1316624.1316632.
4. Mark S. Cary. 1978. The role of gaze in the initiation of conversation. *Social Psychology*: 269-271.
5. Alex Colburn, Michael F. Cohen, and Steven Drucker. 2000. The role of eye gaze in avatar mediated conversational interfaces. *Sketches and Applications*, Siggraph'00.
6. Mark Cook. 1977. Gaze and Mutual Gaze in Social Encounters: How long – and when – we look others "in the eye" is one of the main signals in nonverbal communication. *American Scientist* 65.3: 328-333.
7. Laura Dabbish and Robert E. Kraut. 2004. Controlling interruptions: awareness displays and social motivation for coordination. In *Proceedings of the 2004 ACM conference on Computer supported cooperative work* (CSCW '04). ACM, New York, NY, USA, 182-191. http://dx.doi.org/10.1145/1031607.1031638
8. Daniela Damian, Luis Izquierdo, Janice Singer, and Irwin Kwan. 2007. Awareness in the wild: Why communication breakdowns occur. In *Proceedings of the International Conference on Global Software Engineering* (ICGSE'07), 81-90.
9. Deepak Akkil and Poika Isokoski. 2016. Gaze Augmentation in Egocentric Video Improves Awareness of Intention. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (CHI '16). ACM, New York, NY, USA, 1573-1584. DOI: https://doi.org/10.1145/2858036.2858127
10. Paul Dourish. 1997. Extending awareness beyond synchronous collaboration. *Position Paper for CHI* 97: 249-256.
11. Paul Dourish and Sara Bly. 1992. Portholes: supporting awareness in a distributed work group. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '92), Penny Bauersfeld, John Bennett, and Gene Lynch (Eds.). ACM, New York, NY, USA, 541-547. http://dx.doi.org/10.1145/142750.142982
12. Thomas Erickson and Wendy A. Kellogg. 2000. Social translucence: an approach to designing systems that support social processes. *ACM Trans. Comput.-Hum. Interact.* 7, 1 (March 2000), 59-83. http://dx.doi.org/10.1145/344949.345004
13. Thomas Erickson, David N. Smith, Wendy A. Kellogg, Mark Laff, John T. Richards, and Erin Bradner. 1999. Socially translucent systems: social proxies, persistent conversation, and the design of "babble". In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems* (CHI '99). ACM, New York, NY, USA, 72-79. http://dx.doi.org/10.1145/302979.302997
14. Motahhare Eslami, Aimee Rickman, Kristen Vaccaro, Amirhossein Aleyasen, Andy Vuong, Karrie Karahalios, Kevin Hamilton, and Christian Sandvig. 2015. "I always assumed that I wasn't really that close to [her]": Reasoning about Invisible Algorithms in News Feeds. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (CHI '15). ACM, New York, NY, USA, 153-162. http://dx.doi.org/10.1145/2702123.2702556
15. Danyel Fisher and Paul Dourish. 2004. Social and temporal structures in everyday collaboration. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '04). ACM, New York, NY, USA, 551-558. http://dx.doi.org/10.1145/985692.985762
16. George. W. Furnas. 1986. Generalized fisheye views. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '86), Marilyn Mantei and Peter Orbeton (Eds.). ACM, New York, NY, USA, 16-23. http://dx.doi.org/10.1145/22627.22342
17. Susan R. Fussell, Leslie D. Setlock, and Elizabeth M. Parker. 2003. Where do helpers look?: gaze targets during collaborative physical tasks. In *CHI '03 Extended Abstracts on Human Factors in Computing Systems* (CHI EA '03). ACM, New York, NY, USA, 768-769. http://dx.doi.org/10.1145/765891.765980
18. Susan R. Fussell, Leslie D. Setlock, Jie Yang, Jiazhi Ou, Elizabeth Mauer, and Adam D. I. Kramer. 2004. Gestures over video streams to support remote collaboration on physical tasks. *ACM Trans. Hum-Comput. Interact*. 19, 3 (September 2004), 273-309. http://dx.doi.org/10.1207/s15327051hci1903_3
19. Maia Garau, Mel Slater, Vinoba Vinayagamoorthy, Andrea Brogni, Anthony Steed, and M. Angela Sasse. 2003. The impact of avatar realism and eye gaze control on perceived quality of communication in a shared immersive virtual environment. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '03). ACM, New York, NY,

USA, 529-536. http://dx.doi.org/10.1145/642611.642703

20. Ge Gao, Bin Xu, Dan Cosley, and Susan R. Fussell. 2014. How beliefs about the presence of machine translation impact multilingual collaborations. In *Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing* (CSCW '14). ACM, New York, NY, USA, 1549-1560. http://dx.doi.org/10.1145/2531602.2531702
21. Eric Gilbert. 2012. Designing social translucence over social networks. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '12). ACM, New York, NY, USA, 2731-2740. http://dx.doi.org/10.1145/2207676.2208670
22. Erving Goffman. 1963. *Behavior in Public Places*. New York: Free Press.
23. Erving Goffman. 1984 [originally 1959]. *The Presentation of Self in Everyday Life*. Harmondsworth: Penguin.
24. Charles Goodwin, and John Heritage. 1990. Conversation analysis. *Annual review of anthropology* 19: 283-307.
25. Joy E.Hanna, and Susan E. Brennan. 2007. Speakers' eye gaze disambiguates referring expressions early during face-to-face conversation. *Journal of Memory and Language* 57.4: 596-615.
26. Keita Higuch, Ryo Yonetani, and Yoichi Sato. 2016. Can Eye Help You?: Effects of Visualizing Eye Fixations on Remote Collaboration Scenarios for Physical Tasks. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (CHI '16). ACM, New York, NY, USA, 5180-5190. DOI: https://doi.org/10.1145/2858036.2858438
27. Gordon Hull, Heather Richter Lipford, and Celine Latulipe. 2011. Contextual gaps: Privacy issues on Facebook. *Ethics and Information Technology* 13, 4: 289–302. http://dx.doi.org/10.1007/s10676-010- 9224-8
28. Hiroshi Ishii, and Minoru Kobayashi. 1992. ClearBoard: a seamless medium for shared drawing and conversation with eye contact. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (CHI'92), Penny Bauersfeld, John Bennett, and Gene Lynch (Eds.). ACM, New York, NY, USA, 525-532. http://dx.doi.org/10.1145/142750.142977
29. Jeroen Janssen, Gijsbert Erkens, and Paul A. Kirschner. 2011. Group awareness tools: It's what you do with it that matters. *Computers in Human Behavior*27.3: 1046-1058.
30. Adam Kendon. 1967. Some functions of gaze direction in social interaction. *Acta psychologica 26*: 22-63.
31. Kana Misawa, Yoshio Ishiguro, and Jun Rekimoto. 2012. LiveMask: a telepresence surrogate system with a face-shaped screen for supporting nonverbal communication. In *Proceedings of the International Working Conference on Advanced Visual Interfaces* (AVI '12), Genny Tortora, Stefano Levialdi, and Maurizio Tucci (Eds.). ACM, New York, NY, USA, 394-397. http://dx.doi.org/10.1145/2254556.2254632
32. Kana Misawa, Yoshio Ishiguro, and Jun Rekimoto. 2012. Ma petite chérie: what are you looking at?: a small telepresence system to support remote collaborative work for intimate communication. In *Proceedings of the 3rd Augmented Human International Conference* (AH '12). ACM, New York, NY, USA, Article 17, 5 pages. DOI=http://dx.doi.org/10.1145/2160125.2160142.
33. Andrew F. Monk, and Caroline Gale. 2002. A look is worth a thousand words: Full gaze awareness in video-mediated conversation. *Discourse Processes* 33.3: 257-278.
34. Jens Edlund, Samer Al Moubayed, and Jonas Beskow. 2011. The Mona Lisa gaze effect as an objective metric for perceived cospatiality. In *Proceeding of International Conference on Intelligent Virtual Agents* (IVA'11). Springer Berlin Heidelberg, 439–440.
35. Helen Nissenbaum. 2009. *Privacy in context: Technology, policy, and the integrity of social life*. Stanford University Press.
36. Helen Nissenbaum. 2011. A contextual approach to privacy online. *Daedalus* 140, 4: 32–48.
37. David G. Novick, Brian Hansen, and Karen Ward. 1996. Coordinating turn-taking with gaze. In *Proceedings of Fourth International Conference on Spoken Language* (ICSLP'96). IEEE. 1888-1891.
38. Hiroaki Ogata , Kenji Matsuura, and Yoneo Yano. 1996. Knowledge awareness: Bridging between shared knowledge and collaboration in sharlok. In *Proceedings of Educational Telecommunications*: 232-237.
39. Omnidirectional camera. 2016. Retrieved January 12, 2017 from https://en.wikipedia.org/wiki/Omnidirectional_camera
40. Antti Oulasvirta, Mika Raento, and Sauli Tiitta. 2005. ContextContacts: re-designing SmartPhone's contact book to support mobile awareness and collaboration. In *Proceedings of the 7th international conference on Human computer interaction with mobile devices & services* (MobileHCI '05). ACM, New York, NY, USA, 167-174. http://dx.doi.org/10.1145/1085777.1085805
41. Pan–tilt–zoom camera. 2016. Retrieved January 12, 2017 from

https://en.wikipedia.org/wiki/Pan%E2%80%93tilt%E2%80%93zoom_camera

42. Elin Rønby Pedersen and Tomas Sokoler. 1997. AROMA: abstract representation of presence supporting mutual awareness. In *Proceedings of the ACM SIGCHI Conference on Human factors in computing systems* (CHI '97). ACM, New York, NY, USA, 51-58. http://dx.doi.org/10.1145/258549.258584
43. Christopher Peters, Catherine Pelachaud, Elisabetta Bevacqua, Maurizio Mancini, and Isabella Poggi. 2005. A model of attention and interest using gaze behavior. In *Proceeding of International Workshop on Intelligent Virtual Agents*. Springer Berlin Heidelberg, 229-240.
44. Alex Poole, and Linden J. Ball. 2006. Eye tracking in HCI and usability research. *Encyclopedia of human computer interaction* 1: 211-219.
45. Ramana Rao and Stuart K. Card. 1994. The table lens: merging graphical and symbolic representations in an interactive focus + context visualization for tabular information. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '94), Beth Adelson, Susan Dumais, and Judith Olson (Eds.). ACM, New York, NY, USA, 318-322
46. Ricoh Theta S camera. 2016. Retrieved January 12, 2017 from https://theta360.com/en/about/theta/s.html
47. Markus Rittenbruch, and Gregor McEwan. 2009. An historical reflection of awareness in collaboration. *Awareness Systems*: 3-48.
48. Randy Stein and Susan E. Brennan. 2004. Another person's eye gaze as a cue in solving programming problems. In *Proceedings of the 6th international conference on Multimodal interfaces* (ICMI '04). ACM, New York, NY, USA, 9-15. http://dx.doi.org/10.1145/1027933.1027936
49. John C. Tang, and Ellen Isaacs. 1992. Why do users like video?. *Computer Supported Cooperative Work* 1.3: 163-196.
50. Joseph B. Walther, Brandon Van Der Heide, Artemio Ramirez, Judee K. Burgoon, and Jorge Peña. 2015. *Interpersonal and hyperpersonal dimensions of computer-mediated communication*. Wiley.
51. Bernard Weiner. 2001. Intrapersonal and interpersonal theories of motivation from an attribution perspective. *Student motivation*: 17-30.
52. Suzanne Weisband. 2002. Maintaining awareness in distributed team collaboration: Implications for leadership and performance. *Distributed work*: 311-333.
53. Timo Wolf , Adrian Schröter, Daniela Damian, Lucas D. Panjer, and Thanh HD Nguyen. 2009. Mining task-based social networks to explore collaboration in software teams. *IEEE Software* 26.1: 58-66.
54. Bin Xu, Pamara Chang, Christopher L. Welker, Natalya N. Bazarova, and Dan Cosley. 2016. Automatic Archiving versus Default Deletion: What Snapchat Tells Us About Ephemerality in Design. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing* (CSCW '16). ACM, New York, NY, USA, 1662-1675. http://dx.doi.org/10.1145/2818048.2819948