An Eye For Design: Gaze Visualizations for Remote Collaborative Work

Sarah D'Angelo
Northwestern University
Evanston, IL USA
sdangelo@u.northwestern.edu

Darren Gergle
Northwestern University
Evanston, IL USA
dgergle@northwestern.edu

ABSTRACT
In remote collaboration, gaze visualizations are designed to display where collaborators are looking in a shared visual space. This type of gaze-based intervention can improve coordination, however researchers have yet to fully explore different gaze visualization techniques and develop a deeper understanding of the ways in which features of visualizations may interact with task attributes to influence collaborative performance. There are many ways to visualize characteristics of eye movements, such as a path connecting fixation points or a heat map illustrating fixation duration and coverage. In this study, we designed and evaluated three unique gaze visualizations in a remote search task. Our results suggest that the design of gaze visualizations affects performance, coordination, searching behavior, and perceived utility. Additionally, the degree of task coupling further influences the effect of gaze visualizations on performance and coordination. We then reflect on the value of gaze visualizations for remote work and discuss implications for the design of gaze-based interventions.

ACM Classification Keywords
H.5.3. Information Interfaces and Presentation (e.g. HCI): Group and organizational interfaces - collaborative computing, computer-supported collaborative work

Author Keywords
Eye-Tracking; Design; Remote Collaboration; Gaze Visualization; Dual Eye-Tracking

INTRODUCTION
Remote collaboration is increasingly common and is changing how we work with and learn from others. As an illustration of this growth, the last few years have seen a rapid increase in distance learning programs that serve to broaden access to instructors, course content and learning activities. However, these distributed environments lack many of the rich interpersonal cues that make for effective learning experiences in a co-located classroom environment [3]. For example, the ability to observe students and infer their attentional state happens naturally in co-located environments but, in remote scenarios, the non-verbal cues that provide clues to a learner's attentional allocation are often not visible. A second example can be seen in the workplace, where distributed work teams are becoming commonplace because of the continual rise in globalization. In response to this growth, researchers have continued to develop new video conferencing tools and telepresence systems that aim to improve our ability to meet and effectively interact with remote colleagues. However, many of these systems still lack the ability to display important non-verbal cues that are critical to effective interaction and communication [16, 9].

To address these challenges, researchers have been developing techniques that integrate non-verbal cues into distributed settings in an effort to enhance the remote collaboration experience [7, 13, 8]. One technique that shows considerable promise is the use of gaze visualizations [4, 5, 6, 18, 26, 22]. This involves collecting eye movement data from each person in a pair and displaying that information on a partner's screen. The integration of gaze visualizations into remote work settings aims to improve gaze awareness, i.e. the collaborator's ability to understand what their partner is attending to during a collaborative task. Initial results suggest this method can improve coordination [4, 26, 5]; however, researchers have yet to fully explore different types of gaze visualization and develop a deeper understanding of the ways in which particular features of the gaze visualizations may interact with specific task attributes to influence collaborative performance.

In previous work, gaze is most commonly illustrated as a single visual marker representing an individual's current fixation point [4, 26], or as a path connecting each fixation point by a line representing the saccade [27, 22]. These different visualizations afford different interpretations; for example, a gaze path can reveal spatiotemporal information by connecting sequences of objects or areas of interest in the visual space, whereas a single visual marker cannot. Further, there are many other ways to depict gaze information to support specific tasks. For example, researchers designed an unobtrusive gaze visualization for pair programmers that illustrates where a partner is looking as a rectangle in the right margin of a document [5]. Although this visualization is less spatially precise than traditional methods, results show that the subtle design improved coordination without any of the distracting characteristics (e.g., visual occlusion) shown in previous studies [6]. This suggests that the design of gaze visualizations can be altered to better support coordination for specific tasks.

CHI 2018, April 21–26, 2018, Montreal, QC, Canada

ISBN 978-1-4503-5620-6/18/04. . . $15.00

DOI: https://doi.org/10.1145/3173574.3173923

Figure 1. Example of each gaze visualization.

Thus, there are at least two important factors to consider when developing gaze visualizations to support remote collaborative work. The first is a detailed understanding of the precise form of the gaze visualization: Does it persist over time or only provide an immediate indication of where a partner is looking? Is the visual marker transparent or opaque? Is the display "always on" or do the collaborative partners determine when it is displayed? The second important consideration is a detailed understanding of the particular task attributes: Does the task exhibit or require sequential interdependence? To what extent are the specific physical actions loosely or tightly coupled? And so on. Each of these task attributes impose different requirements on the pair and likely affect the extent to which a given gaze visualization design, and the corresponding awareness it provides, is useful.

To begin to understand these factors, we developed and evaluated three exploratory gaze visualizations for remote collaboration and examined their usefulness across two degrees of task coupling (see Figure 1). Each visualization was designed to represent a different feature of eye movements. The heat map visualization highlights fixation duration and general spatial coverage by marking where someone has looked and darkening the color to show how long they looked there. The shared area representation displays a circle around an area when the pair looks at the same thing together at the same time, to illustrate mutual gaze. Lastly, the path visualization illustrates the current fixation point and saccades by displaying a line connecting the current fixation with a previous fixation.

In this paper we present an experimental study that evaluates three distinct gaze visualizations across the same visual search task with varying degrees of task coupling. Our results show that features of gaze visualizations can impact how pairs coordinate on object locations and search the visual space together. The degree of coupling also influences searching behavior and the effect of displaying gaze visualizations. Further, users perceive the value of the gaze visualizations differently based on specific design features such as how much information is displayed at a given moment. We provide design recommendations for leveraging the features of gaze visualizations and task properties to create more effective gaze visualizations.

BACKGROUND

When pairs collaborate in face-to-face settings they receive a wealth of non-verbal cues from their partner, such as where their partner is looking [19, 10, 23, 24]. Eye movements, in particular, provide valuable information about a partner's allocation of attention, and they have two prominent features: fixations inform us about what people are focusing on and for how long, and saccades reveal how people shift their attention from one fixation point or object to another.

The ability to observe fixations and saccades—or more generally, to establish gaze awareness—provides valuable information about what a person is attending to and it has been shown to affect communication patterns [6] and joint attention [26]. For example, previous studies on collaborative search tasks show that pairs are faster at finding a target when their partner's gaze is displayed because they use the gaze visualization to coordinate on the location of the target [4]. Furthermore, gaze visualizations have been shown to increase learning gains by facilitating coordination between students discussing complex diagrams [26], and serve as an effective referential pointer [2]. However, establishing gaze awareness may not always be beneficial; for example, attending to unintentional eye movements or noisy recordings [12, 21] has been found to be disruptive in tightly-coupled tasks [6].

The conflicting results in the literature around the value of gaze visualizations may result, in part, from the fact that the properties of the task interact with the design chosen for the gaze visualization in a way that affects when gaze awareness is beneficial. One example of this can be seen in task coupling. In a loosely-coupled search task, gaze awareness allows pairs to partition the visual space and effectively search it without needing language to coordinate [4]. Yet, in tightly-coupled tasks that require a lot of back and forth between members of a pair, a continuous display of gaze can be distracting and misleading because gaze information is not always intentional [6]. The variation in task features could account for these different findings. In the first task, gaze awareness supports division of labor, but in the second task unintentional gaze signals disrupt tight coordination. Therefore, it is important to understand how the task features may influence how remote collaborators perceive and use gaze visualizations.

Researchers have investigated a range of different task features [2, 6, 26]; however, they often do so in the context of a single visual representation of gaze for a specific task. As a result, we cannot compare different features of gaze visualization and understand how they interact with task properties. For example, while the most common visualization is a single point or gaze cursor, there are more abstract or subtle ways to visualize gaze information. For example, an overlay of a partner's head can provide limited information about where they are looking [11]. Additionally, gaze information has been displayed in the periphery to subtly attract attention to specific areas on the screen [1] or as a spotlight to restrict view to a specific area [14, 15]. While these findings suggest that the design of gaze representations can facilitate coordination in different ways, we have yet to evaluate different designs across the same task to understand how visual representations can be designed to best support specific task requirements.

Recent work on co-located collaboration has explored different gaze visualizations to support collaborative search [31]. This work and related work in remote collaboration [6], reveal that while gaze awareness can support communication, the visual display can be distracting and disrupt coordination. However as previously noted, there are many features of gaze that can visualized. Thus, differing approaches to gaze visualizations make it difficult to compare results between studies and systems, since it's not known how the design of the visualization alters its effectiveness. A comparative evaluation of different gaze visualization techniques can ensure that other work utilizing gaze visualization maximizes potential benefit. To effectively integrate gaze information into remote work, we need to think critically about the visual representations of gaze information and how it relates to the task requirements of the pair. The current literature has tended to focus on a single gaze visualization applied to a single task, leaving open the question of what differences are caused by the gaze visualization, the task properties, or an interaction of the two. In this work, we address this question by varying the visualization technique and task features in a controlled experimental setting.

THE CURRENT STUDY

We investigate how the design of the gaze visualization and the particular requirements of the task interact to affect communication processes and ultimately influence collaborative performance. We design and implement three different gaze visualization techniques [1](see the accompanying video for details) and then examine pair performance using each technique at two levels of task-coupling during a remote visual search task (see Figure 2). Each of the three different visualizations are designed to highlight different collaborative gaze features. The visualizations include: a *heat map*, a *shared area* visualization that is displayed when pairs look at the same area at the same time, and a *path* representation. We also include a *no visualization* baseline (see Figure 1). For each visualization, pairs search for objects collaboratively and independently.

Gaze Visualizations

The design of the following gaze visualizations resulted from iterative testing and early user feedback. It's important to note that the designs are intended to be exploratory and they aim to illustrate various visualization *attributes* that are important to consider when designing systems to support remote collaboration as opposed to defining optimal visualization designs.

Heat map: Heat map visualizations are a common way to depict eye movement patterns as they accrue over time and as a result they highlight general areas of interest within a scene. As a real-time visualization, the heat map provides information on where participants have searched which can support effective division of the work space and help partners avoid searching the same areas. Our implementation of the heat map visualization shows where a partner has been looking in yellow; the color darkens to orange and then to red as the partner revisits the same location or fixates in the same location for an extended period. Each fixation is displayed as a circle (80 px diameter) at 70% opacity. After 20 fixations that overlap within a window of fourteen seconds the color transitions from yellow, to orange, followed by dark orange, red, and lastly dark red. The color accumulates as the partner looks around the visual space. However, previous fixation points will fade out after fourteen seconds to help avoid occluding the entire work space.

Shared area: The shared area visualization is not displayed continuously; instead it is designed to highlight moments when the partners are looking at the same place together. We expect this may support coordination between the collaborators while avoiding distractions of a continuous display. Pairs can use the visualization to facilitate searching together while coordinating on the object location. The shared area visualization displays a gray circle outline (5pt thick, 50px diameter) when the participants look at the same area at the same time. This is defined as any time when the straight line distance between both gaze coordinates is less than 125px or approximately one inch of the work space. The calculation is performed in real time by our gaze sharing system and the visualization is automatically displayed when the threshold is achieved. The center of the circle is displayed at the midpoint between the gaze coordinates. As the pair continues to fixate on the area together, the circle continues to be visible until one person looks outside the shared area for more than 100ms.

Path: The path visualization is designed to highlight connections between objects in the shared visual space as well as display a continuous real-time representation of where a collaborator is currently looking. The trail of gaze information (i.e., the amount of time the previous fixation is visible on screen) is three seconds. This visualization is designed to support efficient referential communication between participants by using the gaze cursor as a pointer in the shared visual space. Therefore, we expect pairs to use the visualization to effectively refer to object locations. However, the "always on" nature of the display may distract pairs or cause them to follow their partner's gaze cursor. This visualization displays the participant's current fixation point illustrated as a black circle (10px diameter) and a connection to the previous fixation point illustrated as a line (1pt thick) to a semi-transparent red circle (10px diameter, 10% opacity). A previous fixation is only displayed if the fixation duration was longer than 700ms. When the previous fixation is not displayed, participants see the current fixation point without a path.

[1]We make this code available so that other researchers can see the implementation details for our various visualizations as well as use the software to develop their own. http://collablab.northwestern.edu/EyeTracking.php

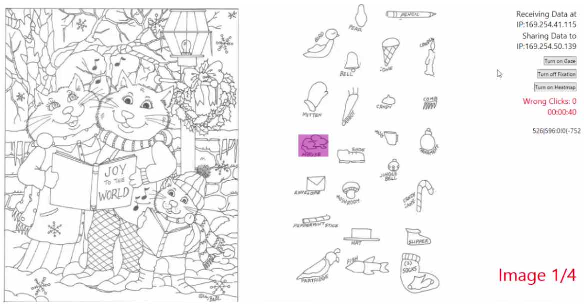

Figure 2. Screenshot of the search task for collaborative search (as viewed by one partner). The target object is highlighted in purple.

Task Coupling

In *collaborative* searches, pairs search for the same object together and cannot move on to the next object until both participants locate the object. Participants are instructed by a researcher to help each other locate the object. In the *independent* searches participants search for different objects. Once they locate their object they immediately move on to the next object. Participants are informed by a researcher that they are looking for different objects and only need to find their respective objects to move on.

METHOD

Participants

Ninety-six college students participated in the experiment (48 pairs). Seventy-six percent of participants were female. Participants ranged in age from 18 to 25 with 75% in the 18-21 range and 25% in the 22-25 range. Forty-two percent of participants were Caucasian (34% Asian/Pacific Islander, 4% African-American, 7% Hispanic, and 13% Mixed Race / Other). Pairs had no past experience with gaze visualizations. All participants received $10 compensation for their participation. In this paper, participants are identified by pair number and computer ID (e.g. Pair 1 Computer A is P1CA).

Apparatus

The experiment setup consists of two Tobii 4C remote eye trackers and two 15" Lenovo laptops with wireless mice. A trifold is used as a visual barrier to prevent the pair from seeing each other while allowing them to communicate verbally. The eye trackers capture the gaze patterns of each participant and our software sends the coordinates to the partner's display to be visually represented. The task is locally networked to mirror all actions on each display. The computers are locally networked to minimize any potential lag or delays and to display visualizations in real time. Eye movements are sampled at a rate of 30Hz and participants complete a 5-point calibration to an accuracy between .5 and 1 degree of visual angle.

Experimental Design

We employ a 2×4 experimental design with task coupling (collaborative, independent) and gaze visualization (heat map, shared area, path, no visualization) as within-subject factors. There are four unique images that were pre-tested to control for difficulty. The image order is fixed and participants always complete collaborative searches before performing independent searches for each image. For each unique task coupling and gaze visualization condition, participants saw one of four images. Each image occurred equally often across participants and conditions. We fully counterbalanced the order of gaze visualizations, giving 24 unique orders. For each pair, the order that was used for the collaborative search was repeated in the independent search.

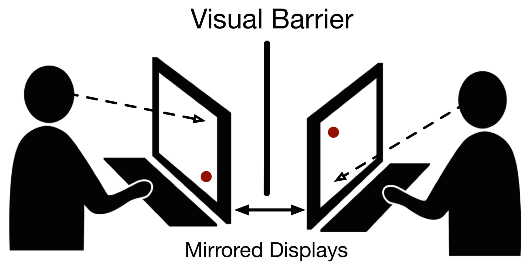

Figure 3. Experiment setup.

Procedure

When participants arrive at the lab, they are directed to the experiment room and provided with written consent materials. After the participants provide their written consent the researcher explains that they will be looking for objects in an image and that sometimes they will work together to find the same object and other times they will have to find different objects by themselves. They are told that the entire experiment will last approximately 30 minutes. They are also told that they will be able to see where their partner is looking on the screen during the task, and that the way their partner's gaze information is displayed will change throughout the experiment.

At the start of the experiment, the eye trackers are calibrated for the participants (5-point calibration). Next, the participants complete the search tasks. For each of the gaze visualization techniques and the no visualization baseline, participants are shown an image with objects hidden inside and must find six objects in the image. On average, participants completed a single search task in 4.3 minutes. Half of the objects are found collaboratively, and half independently (see Figure 2).

When searching collaboratively, both members of the pair must locate the same object. Once one participant finds the object they must help their partner find the object before they can both move on to the next object. Each participant must click on the object as soon as they discover it to signify that they have located it. When searching independently, the participants are looking for different objects. Once a participant finds their object, they click on it and move on to the next object, independent of when their partner locates their object.

Pairs are instructed to find the objects as quickly as possible. Elapsed time and incorrect clicks are displayed on screen to encourage the pairs to find the objects as fast as they can without making incorrect guesses. After each search task, the participants fill out a survey about the gaze visualization they just saw (excluding the no visualization condition). Following the survey, they continue to the next search task. The duration of the experimental task is approximately 20 minutes. For the remaining 10 minutes, participants are interviewed about their perceptions of the gaze visualizations.

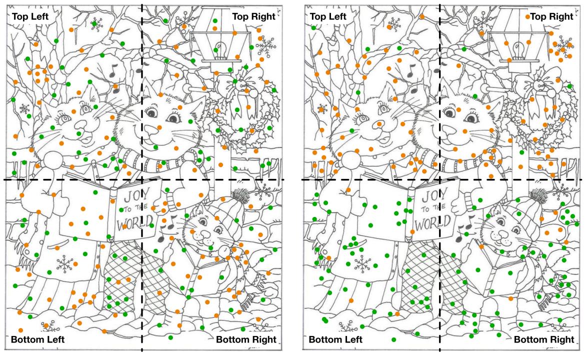

Figure 4. Example of high (left) and low (right) Quadrant overlap. In the high overlap the orange and green fixation indicators display a great deal of overlap within the quadrants; while in the low overlap panel the pairs segment the space with one person focusing on the top (orange) and the other on the bottom (green).

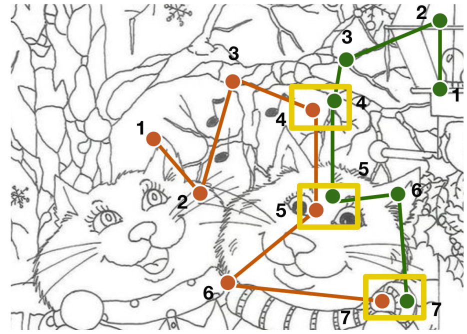

Figure 5. Example of gaze coordinates overlapping in time (highlighted in yellow).

Measures

Our evaluation of the gaze visualizations includes both outcome and process measures as well as self-reported data. Outcome measures include completion time and coordination time which contribute to understanding how the design of gaze visualizations impact task performance. Process measures include search patterns and content analysis which provide a closer look at what contributed to changes in performance. Lastly, self-reported data includes post-task surveys and interviews which allow us to understand participants' perceived differences between the visualizations. We examine multiple outcome variables and measures in a mixed methods approach in order to best illustrate how different visualization techniques can impact coordination at various levels.

Our approach for analysis of outcome measures and searching behavior applies linear mixed models (LMM). This technique has several advantages over ANOVA (which is a special case of LMM with only fixed effects) for example, it accounts for both fixed and random effects and adjusts standard error (SE) to better accommodate for repeated measures [30]. We would like to note that the degrees of freedom (DFs) differ because of the use of REML (in JMP) which adjusts DFs to best account for linear correlations and correlated errors in the model [25]. Additionally, we use the log transformed time to better approximate a normal distribution and correct for skew in the data that is commonly found in performance time data.

Completion Time: The completion time is measured in seconds from when the search object is displayed to when the object is clicked on by the participant. This measure differs depending on the coupling of the pair. In the collaborative condition, the overall completion time is recorded when the last participant of the pair locates the object (see Figure 6). While in the independent condition, completion time is recorded separately as the time taken for each participant to find the object. Therefore, there is no analysis of main effect for collaborative vs. independent time because they are different units of analysis: collaborative time is recorded at the pair level whereas independent times are recorded for each participant in a pair. This measure reflects overall performance on the task.

Coordination Time: In the collaborative search, the time between when the first participant locates the object and when their partner finds the object is measured in seconds (see Figure 6). This measure reflects the coordination between the pair or how long it takes for one participant to successfully describe the location of the object to the other participant.

Our general approach for analysis of completion time and coordination time applies a linear mixed model with gaze visualization (path, heat map, shared area, no visualization) and image order (1-4) as within-pair factors. Collaborative measures of completion and coordination time are modeled at the pair level while independent completion time is modeled for each participant in a pair. We therefore model the covariance structure and include the participant or pair as a random effect to account for the fact that each performed in every condition and as a result their observations are not independent.

Search Patterns: Eye movement data is recorded and analyzed to capture how the pairs search the visual space. The extent to which pairs search the same area is measured in the *total gaze overlap* per quadrant of the image [4]. In other words, we calculate the fraction of time during which participants looked in the same quadrant as their partner for each object (see Figure 4). This measure reflects, on average, how often participants look at the same quadrant as their partners while searching for an object. When pairs successfully divide the visual space we would expect to see low quadrant overlap, while pairs who repeatedly search the same areas as their partner would have higher quadrant overlap.

We also measure *gaze overlap in time* as the proportion of time pairs spend looking in the same area at the same time [5] (see Figure 5). While the previous measure looks more broadly at how pairs divide the visual space over the course of a searching period, this measure only captures when pairs look at the same point at the same time. We consider pairs to be looking at the same point if the distance between their gaze coordinates is less than 50px (approximately half an inch). This measure

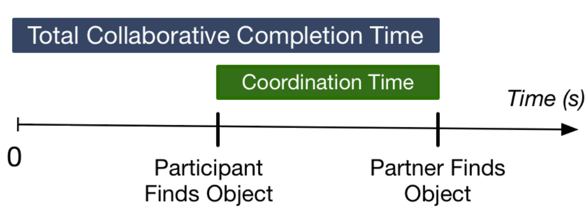

Figure 6. Collaborative completion time and coordination time.

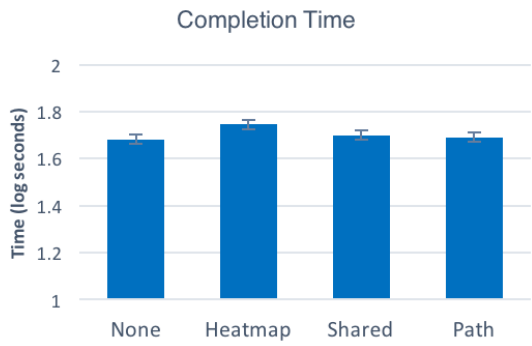

Figure 7. Overall completion time by gaze visualization (in log seconds). Error bars indicate SEs.

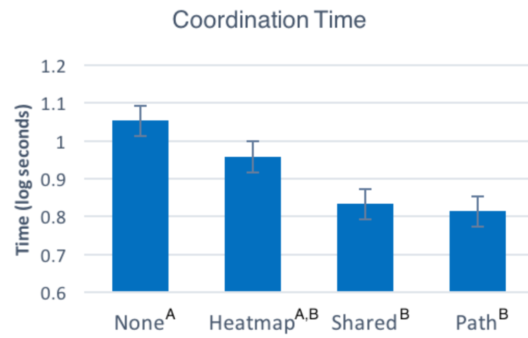

Figure 8. Overall coordination time by gaze visualization condition (in log seconds). Error bars indicate SEs (levels that do not share the same letter are significantly different, based on Tukey's HSD test).

captures how often the pairs were looking at their partner's gaze visualization while searching for an object. Participants who frequently look at their partner's gaze visualization would have high gaze overlap. We expect pairs who used the gaze visualization to coordinate on object locations to have high overlap in time while pairs who were searching independently to have low overlap in time.

Our analysis of search patterns applies a linear mixed model regression with gaze visualization (path, heat map, shared area, and no visualization), image order (1-4), and coupling (collaborative and independent) as within-pair factors. Participant pair is modeled as a random effect.

Content Analysis: Conversations were transcribed and we analyzed moments when pairs were communicating about the location of an object. We specifically focus on how pairs use language to reference the object location. For example, using deictic references (e.g. "here") or relative location descriptions (e.g. "to the left of the cat"). We compare how pairs communicate about the location of the object with the different gaze visualizations compared to without a gaze visualization.

Survey and Interviews: The visualization survey asked participants to report on how they felt about the effectiveness of each of the gaze visualization techniques. Participants completed a five question survey after each visualization condition. Participants are asked to record the extent to which they agree on a 5-point Likert scale (ranging from "1 - Strongly Disagree" to "5 - Strongly Agree") see Figure 12 for statements.

At the conclusion of the experiment, participants reported their preference each gaze visualization (on a scale of 1-3). The researcher also conducted semi-structured interviews about the participants' perceived value of the visualizations. Responses were transcribed and reviewed by the research team to identify common themes. Then representative quotes were identified to reflect how features of the gaze visualizations improved or disrupted their ability to communicate with their partner.

RESULTS

Completion Time

Completion time (log seconds) results for the collaborative searches indicate that the heat map visualization (M(SE) = 1.745(0.028)) was the only gaze condition that was different than the condition group mean (M(SE) = 1.706(0.013); t = 2.20, $p = 0.029$), demonstrating that the pairs were on average slower to find objects with the heat map visualization (see Figure 7). This suggests that the heat map visualization may have been more disruptive compared to the other visualizations and no visualization. Interestingly, when we look at the completion time for the independent searches (M(SE) = 1.378(0.016)) we do not see any differences across conditions. This suggests that the negative effects of the heat map may only impact performance when pairs are searching together and using the gaze visualization to coordinate on object locations.

Coordination Time

Coordination time (log seconds), or the amount of time it takes one person to help their partner find the object, results show a significant main effect of gaze visualization, ($F(3,124.8) = 5.85$, $p = 0.0009$ see Figure 8). A Tukey's HSD test shows pairs are significantly faster at coordinating on the location of an object with the path visualization (M(SE) = 0.82(0.05); $p = 0.0008$) and the shared area visualization (M(SE) = 0.082(0.06); $p = 0.0081$) compared to no visualization (M(SE) = 1.07(0.05)). The heat map visualization (M(SE) = 0.94(0.06)) lies somewhere between no visualization, shared area, and path. This suggests that pairs use the gaze visualizations to help communicate about the location of objects – a point we return to in the content analysis.

Searching Behavior

We calculate the proportion of time the pair looked in the same areas while searching for an object as the overall amount of overlap in each quadrant (see Figure 4). We see that pairs spend significantly more time searching the same quadrants in the collaborative searches (M(SE) = 0.73(0.01)) compared to independent searches (M(SE) = 0.64(0.01); $F(1,283.8) = 78.19$, $p < .0001$). We did not see any differences in gaze distribution overlap among the visualizations when the pairs were performing independent searches ($F(1,287.6) = 0.74$, $p = 0.39$). However, we see that when the path visualization is displayed in collaborative searches (M(SE) = 0.78(0.02)), pairs exhibit significantly more quadrant overlap compared to shared area (M(SE) = 0.73(0.02)) and no visualization (M(SE) = 0.72(0.02); $F(1,288.4) = 5.79$, $p = 0.018$), while heat map (M(SE) = 0.75(0.02)) lies in the middle (see Figure 9). The

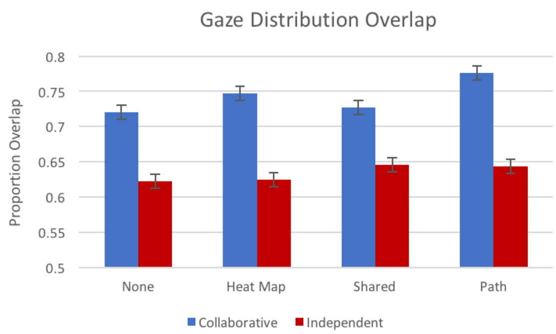

Figure 9. Proportion overall gaze distribution overlap with SE.

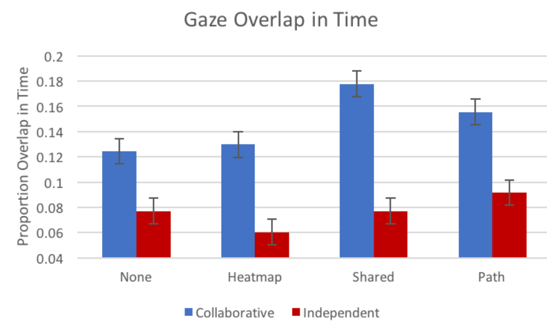

Figure 10. Proportion of gaze overlap in time with SE.

high gaze overlap in the collaborative searches with the path visualization may be due to the distracting characteristics of the "always on" visualization which can cause participants to follow their partner's gaze. We do not see any effect of visualization on the independent searches.

Our measure of overall quadrant overlap captures how pairs divided the visual space to reveal when pairs searched the same quadrant. However, it does not tell us if pairs were looking at the same point at the same time. To understand if pairs searched together in time, we calculated the proportion of time that the pairs are looking in the same 50px radius (see Figure 5). As expected, pairs spend significantly more time searching together in the collaborative searches (M(SE) = 0.15(0.01)) compared to independent searches (M(SE) = 0.08(0.01); F(1,308) = 58.97, $p < .0001$). Showing that when pairs are working together they spend more time searching concurrently. A student's t-test reveals that there is no effect of visualization in independent searches. As expected, pairs are not likely to look at their partner's gaze visualization when searching independently.

In the collaborative searches, we expect to see the most overlap in time when the shared visualization is displayed because the visualization is designed to encourage pairs to look together. To activate the visualization pairs must look in the same area at the same time. Therefore they are likely to stay coupled in time while coordinating on object locations. In contrast, the heat map visualization discourages looking together because the visualization technique occludes the part of image where the collaborator is looking. As expected, we see significantly more overlap in time when the shared area visualization is displayed (M(SE) = 0.18(0.01)) compared to heat map (M(SE) = 0.13 (0.01); F(1,308) = 6.87, $p = 0.009$) and no visualization (M(SE) = 0.13(0.01); F(1,308) = 7.69, $p = 0.006$). We also expect to see high overlap in time with the path visualization because the "always on" feature encourages pairs to attend to their partner's visualization. While this is reflected in the ordering of conditions (see Figure 10), the overlap in the path condition lies in the middle (M(SE) = 0.15(0.01)) and is not significantly different from the other conditions. The impact of visualization technique on searching behavior suggests that gaze visualizations can be designed to help support or discourage looking together.

Content Analysis

When we take a closer look at the coordination period in the transcripts and video recordings we see how the pairs use language and the gaze visualization to help describe and guide their partner to the location of the object. Figure 11 illustrates the common patterns of coordination observed in transcripts and video recordings. For example, with the path visualization participants make statements such as *"I found it, it's right here!"*, while looking at the object. Their partner is then able to see where they are looking and find the hidden object quickly. As depicted in Figure 11, once a participant locates the object they are able to communicate efficiently without need for clarification. Consistent with prior work [6, 2], participants make use of the path visualization as a referential pointer to help coordinate on object locations.

However, when there is no visualization present, the pairs have to spend more time and effort describing the location of the hidden object to their partner (see Figure 11). For example, the following is an exchange from a pair attempting to find an object hidden in a tree branch.

> P14CB: "it's like in the main tree, on the branch, if you go follow the branch and then go to the right branch and then go to the left branch"
>
> P14CA: "wait okay main tree... left branch"
>
> P14CB: "uh huh"
>
> P14CA: "and then go to the right branch?"
>
> P14CB: "yea and then go left again...so like your on the main tree"
>
> P14CA: "uh huh"
>
> P14CB: "turn left"
>
> P14CA: "okay... I turned left"
>
> P14CB: "now keep going"

Looking at the process to find the same object with the shared area visualization, we see decreased conversational effort once the pairs locate the general region where their partner is looking and the shared area visualization is displayed. We see participants using landmarks in the image (see below) or location cues such as "upper right" in order to roughly identify the spatial region of interest. Pairs then quickly scan the landmark or region until the visualization appears and then they use it as a referential pointer (see Figure 11).

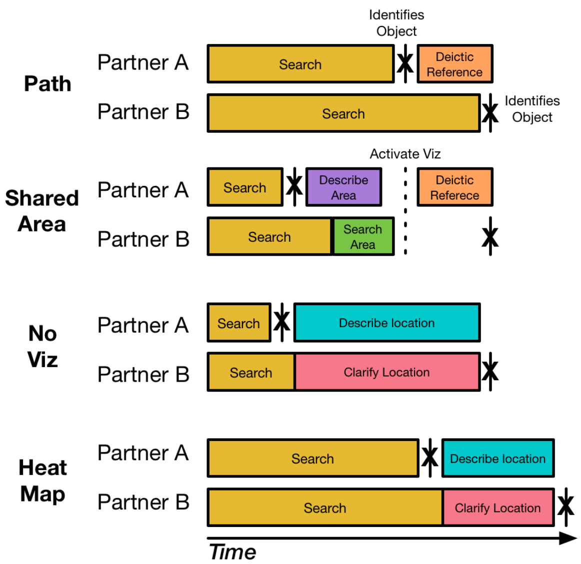

Figure 11. Graphical representation of common coordination processes for each visualization (time duration reflects average time).

P13CA: "I think I found it, see where he is holding the snow flake in his left hand"

[Shared area visualization appears]

P13CA: "right on top of that"

P13CB: "oh yea"

In comparison, the always on heat map visualization allows participants to use the visualization to signal to their partner where they are looking by looking in the same area for an extended period of time. However, the visualization displays the previous 14 seconds of fixations, which can cause confusion with multiple points being displayed at the same time. Continued staring will darken the color of the visualization to help participants clarify the signal.

P8CA: "I found it, it's the guys left"

P8CB: "which guy"

P8CA: "the dogs left foot, leg"

[Heap map visualization darkens]

P8CB: "oh yep yep, nice"

However, if they fixate in that region too long the heat map coloring starts obscure the object underneath the gaze visualization – which led to coordination problems for some pairs.

P45CB: "Just stare at it. Is it where it is turning ... wait. Oh, it's just a big red blob. Just tell me where it is."

The differences in how pairs describe the location of the object across the different gaze visualizations and no visualization baseline are consistent with the coordination time results. We see that pairs require less descriptive language and spend less time coordinating with the path visualization. The shared area visualization requires pairs to initially use descriptive language to activate the visualization and then, once it is activated, they can quickly coordinate using the visualization. Additionally, the heat map visualization can be used to circumscribe the

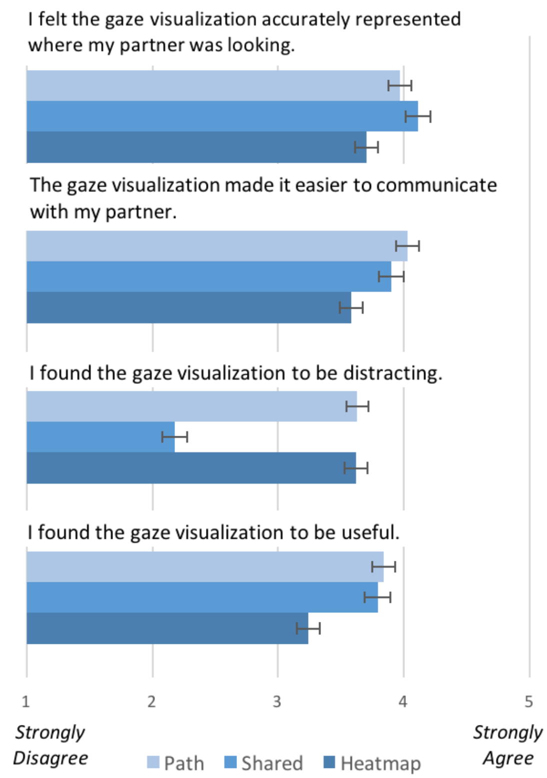

Figure 12. Survey results.

referential domain [17] by marking areas that participants have searched. However, when too much of the visual space is highlighted, it can require more coordination to disambiguate references. With no visualization, the pairs rely entirely on language, which requires more descriptive language use to coordinate on object locations.

Survey Results

Participants reported mixed results on the value of each of the gaze visualizations (see Figure 12). A Wilcoxon Signed Rank comparison for each pair reveals that participants found the heat map visualization to be significantly less useful than the path visualization ($Z = 3.94$, $p < .0001$) and the shared area visualization ($Z = 3.66$, $p = .0003$). There was no significant difference for usefulness between shared area and path. However, participants perceived both path ($Z = 7.90$, $p < .0001$) and heat map ($Z = 7.55$, $p < .0001$) to be significantly more distracting than the shared area visualization. There was no significant difference between path and heat map (see Figure 12). These results indicate that while the path visualization was perceived as more useful than the heat map visualization, it was also perceived as distracting. The results for the shared area visualization show that it was perceived as being more useful without being distracting. A possible explanation for these differences is that the continuous display of the path and heat map visualizations was more distracting than the momentarily displayed shared area visualization.

We see that participants perceived the path ($Z = 3.50$, $p = .0005$) and shared area ($Z = 2.33$, $p = .019$) visualizations as facilitating communication more effectively than the heat

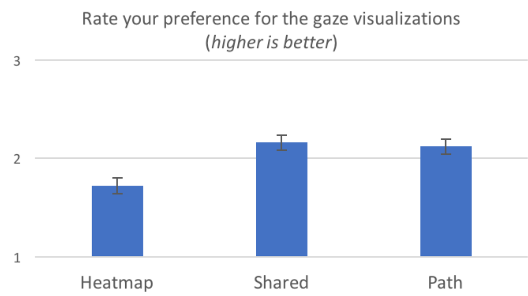

Figure 13. Visualization rating results.

map visualization. There was no difference between path and shared area. As corroborated in the transcripts, participants were able to use the path and shared area visualizations as a referential point to quickly communication about the location of the object. Additionally, participants perceived the path (Z = 2.53, p = .011) and shared area (Z = 3.29, p = .001) visualizations to be more accurate than the heat map visualization. And once again there was no detected difference between the path and shared area visualizations. This may be due to the precision of the visualization. The path and shared area visualizations display a more precise representation of where the participant is looking while the heat map can be ambiguous. All pairs reported high levels of understanding the visualizations (M(SD) = 4.36(0.59)).

When pairs rated their preference for each visualization we see that the path (Z = 2.85, p = .004) and shared area (Z = 3.50, p = .0005) visualizations are preferred more than the heat map visualization. There was no difference between the path and shared area visualizations for overall preference (see Figure 13). A possible explanation for this preference is that while the path visualization was perceived as more distracting than the shared area, we see similar levels of usefulness which may be due to the ability to use it as a referential pointer. Further, pairs perceived the path and the shared area visualization to be better at facilitating communication and more accurate than the heat map visualization. Therefore, pairs may prioritize these task benefits and accept distracting features.

Interviews

In interviews, the participants expressed mixed opinions regarding the value of each of the gaze visualizations. Consistent with the survey results we see a stronger preference for the path and shared area visualizations over the heat map visualization. One feature of the visualizations which contributed to user preference was whether it was always displayed or partially displayed. For example, the path visualization was always on, which contributed to some of its distracting characteristics.

> "I couldn't help but follow it" - P14CB

While the continuous display of the path visualization was distracting, participants expressed that its precision and availability made it useful for coordinating quickly when their partner located the target object.

> "It was the easiest way you could tell where they are looking. It was like a mouse, which was cool. You could just follow it, you weren't really able to follow either of the other two" - P6CB

In contrast, the partially available shared area visualization was perceived to be less distracting than the path visualization, but equally useful. Although it required slightly more effort to activate the display, the lack of continuous information made it less distracting for pairs.

> "With shared gaze we were able to do our own thing but just target the objects together" - P11CA

Furthermore, participants used visual confirmation that pairs were looking in the same place at the same time to coordinate.

> "I think it was easier when you're like, look there, and then you're like, yeah, that's right, because you knew they were looking at the right spot" - P5CB

The heat map visualization was least preferred by participants due to the distracting aspects of a continuous and cumulative display. Heat map visualizations display past fixations (in this study for fourteen seconds) which displays a lot of information. We see that this amount of gaze information being displayed was disruptive to participants.

> "So the heat map I put as the least useful because it stayed for a long time and it was distracting" - P8CA

Consistent with prior work [27], we see that in real time collaboration displaying prior gaze information for extended periods of time can be disruptive. This suggests that while heat map visualizations are a common way to visualize cumulative gaze information, they are less useful for real time collaboration.

DISCUSSION

While remote collaboration continues to become more common, we still face challenges conveying non-verbal cues. We explore how to represent gaze information in distributed environments to facilitate communication between remote collaborators. We see that the design of gaze visualizations and the properties of the task influence how pairs coordinate and allocate attention. Based on the results of this study, we identify features of gaze visualizations that influence coordination. Further, we see that properties of the task such as degree of coupling also affect how gaze visualizations are attended to or ignored. These results have implications for the design of gaze visualizations in many remote tasks including online learning [26, 27], medical education [28], programming [5, 29], trip planning [22], game play [20], and problem solving [2].

Features of Gaze Visualizations

The design of gaze visualizations influences how pairs coordinate. Consistent with prior work, we see that direct representations of gaze such as the path visualization are used as referential pointers [6, 2]. The always on feature of this visualization allowed pairs to coordinate quickly on object locations. While participants perceived this feature to be useful they also reported that it was distracting. We see this reflected in the searching behavior. Pairs spend more time searching together and are more likely to revisit the same areas as their partner with the path visualization. This suggests that they are attending to and following their partner's visualization.

These mixed responses and differences in coordination behavior suggest that there are both advantages and disadvantages to a continuously displayed gaze visualization with respect to remote collaborative tasks.

With the shared area visualization we see an opportunity to achieve the same coordination improvement as well as perceived utility without the distracting effects of a continuously displayed visualization. Pairs were able to successfully activate the shared area visualization and use it to facilitate communication about object locations. We see that pairs have less overall gaze overlap compared to the path visualization but equivalent overlap in time which suggests they are able to search separate areas and effectively come together when locating an object. This form of gaze visualization could be particularly useful in collaborative tasks that require moments of tight coordination in addition to effective division of labor. For example, students studying together would benefit from less distraction while still having the ability to quickly coordinate when they have questions. While this style of visualization is new, we suggest further exploration into the design of partially available gaze visualizations.

The heat map visualization was perceived to be the least useful as well as most distracting. Further, pairs were slower to find objects when the heat map visualization was displayed. However, we do see some improvements in their ability to communicate about the location of the object compared to no visualization. In contrast to the previous visualizations, the heat map discouraged pairs from searching together in time because the visualization occludes the image.

Features of Collaborative Tasks

The degree of coupling between the pair is also an important consideration when incorporating gaze visualizations in remote collaborative work. While the features of the gaze visualizations affect how pairs coordinate in collaborative searches, we see no effect of displaying gaze visualizations in independent tasks. This suggests that pairs are able to ignore gaze visualizations when they are not collaborating. Furthermore, as we would expect for an independent task, searching behavior shows that there is less quadrant overlap and less time spend looking together in independent searches compared to collaborative searches. These results suggest that when pairs believe there is value to knowing where their partner is looking they are more likely to attend to that information. Whereas, when pairs know their partner is searching for a different object they are less likely to attend to their gaze information regardless of how distracting they perceive it to be. Therefore, the design of gaze visualization is especially important for tightly coupled collaborative work while loosely coupled collaboration is not impacted to same degree.

The complexity of the task can also influence what characteristics of coordination to encourage with gaze visualizations. For example, compared to simple collaborative search tasks, the hidden image task presents the opportunity for a person to look in the correct area without correctly identifying the hidden object. Therefore, pairs may have been more likely to search the same areas as their partner when they could not find the object. In other words, simply scanning the image may not result in successfully locating the target. This is in contrast to prior work [4] which illustrated effective division of labor in a simple collaborative search task when a gaze visualization is displayed. The hidden image task may benefit from more gaze overlap. Therefore, the heat map visualizations could be useful in simple collaborative search tasks where participants should not search the same area as their partner. On the other hand, the hidden image task benefits from the shared area visualization which supports some search overlap without the distracting aspects of the path visualization.

Limitations

This work was conducted in a controlled lab environment with a specific focus on collaborative search. This allows us to investigate the specific effect of visualization design on how pairs communicate about locations and objects in a shared visual space. However, it limits our ability to make claims about other kinds of tasks and real world environments. Therefore, we encourage future work to investigate the design of gaze visualizations in other configurations and environments.

Further, we use affordable commercial remote eye trackers to simulate a natural environment in which participants are not physically restricted. However, we acknowledge that this limits the eye tracking quality [12, 21]. Our intention is not to achieve precision but instead we aim to evaluate gaze visualizations in a representative setting for casual eye tracking on personal devices. This allows us to understand how different approaches may fare better or worse with dropped frames or noisy input. For example, the shared area and heat map visualizations are more robust to dropped coordinates and noise compared to the path visualization.

CONCLUSION

In this work, we evaluated three unique gaze visualizations across the same visual search task with varying degrees of task coupling. The results of this study demonstrated that the design of gaze visualizations play a critical role in how they are used to support coordination. Furthermore, the properties of the task determined how the gaze visualizations impact attention. We have identified availability of the visualization and amount of information displayed as important features of visualizations that influence how pairs allocate attention and coordinate with each other. The properties of the task can determine which features are most appropriate to support effective collaboration. As remote collaboration continues to grow in popularity and find new applications it is important to consider the role of non-verbal cues in facilitating communication. Gaze visualizations can be an effective tool for enhancing communication in a variety of contexts. A broader understanding of the features of gaze visualizations can help designers adapt to new contexts quickly and effectively use gaze visualizations to support specific task goals.

ACKNOWLEDGMENTS

We would like to thank Nancy Yao and Jeff Brewer for their contributions to this work as well as acknowledge support from the Google PhD Fellowship in HCI for the first author.

REFERENCES

1. Reynold Bailey, Ann McNamara, Nisha Sudarsanam, and Cindy Grimm. 2009. Subtle gaze direction. *ACM Transactions on Graphics (TOG)* 28, 4 (2009), 100.
2. Ellen Gurman Bard, Robin L Hill, Mary Ellen Foster, and Manabu Arai. 2014. Tuning accessibility of referring expressions in situated dialogue. *Language, Cognition and Neuroscience* 29, 8 (2014), 928–949.
3. Kristen Betts. 2009. Lost in translation: Importance of effective communication in online education. *Online Journal of Distance Learning Administration* 12, 2 (2009).
4. Susan E Brennan, Xin Chen, Christopher A Dickinson, Mark B Neider, and Gregory J Zelinsky. 2008. Coordinating cognition: The costs and benefits of shared gaze during collaborative search. *Cognition* 106, 3 (2008), 1465–1477.
5. Sarah D'Angelo and Andrew Begel. Improving Communication with Shared Gaze Awareness in Remote Pair Programming. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (2017). ACM.
6. Sarah D'Angelo and Darren Gergle. Gazed and Confused: Understanding and Designing Shared Gaze for Remote Collaboration. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (2016). ACM, 2492–2496.
7. Susan R Fussell, Leslie D Setlock, Jie Yang, Jiazhi Ou, Elizabeth Mauer, and Adam DI Kramer. 2004. Gestures over video streams to support remote collaboration on physical tasks. *Human-Computer Interaction* 19, 3 (2004), 273–309.
8. Darren Gergle, Robert E Kraut, and Susan R Fussell. 2013. Using visual information for grounding and awareness in collaborative tasks. *Human–Computer Interaction* 28, 1 (2013), 1–39.
9. David M Grayson and Andrew F Monk. 2003. Are you looking at me? Eye contact and desktop video conferencing. *ACM Transactions on Computer-Human Interaction (TOCHI)* 10, 3 (2003), 221–243.
10. Zenzi M Griffin and Kathryn Bock. 2000. What the eyes say about speaking. *Psychological science* 11, 4 (2000), 274–279.
11. Karl Gyllstrom and David Stotts. 2005. Facetop: Integrated semi-transparent video for enhanced natural pointing in shared screen collaboration. *May* 15 (2005), 1–10.
12. Kenneth Holmqvist, Marcus Nyström, and Fiona Mulvey. 2012. Eye tracker data quality: what it is and how to measure it. In *Proceedings of the symposium on eye tracking research and applications*. ACM, 45–52.
13. 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*. ACM, 525–532.
14. Halszka Jarodzka, Katharina Scheiter, Peter Gerjets, and Tamara Van Gog. In the eyes of the beholder: How experts and novices interpret dynamic stimuli. In *Learning and Instruction* (2010), Vol. 20. 146–154.
15. Halszka Jarodzka, Tamara van Gog, Michael Dorr, Katharina Scheiter, and Peter Gerjets. 2013. Learning to see: Guiding students' attention via a model's eye movements fosters learning. *Learning and Instruction* 25 (2013), 62–70.
16. Steven Johnson, Irene Rae, Bilge Mutlu, and Leila Takayama. 2015. Can you see me now?: How field of view affects collaboration in robotic telepresence. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. ACM, 2397–2406.
17. Robert E. Kraut, Darren Gergle, and Susan R. Fussell. The use of visual information in shared visual spaces: Informing the development of virtual co-presence. In *Proceedings of the 2002 ACM conference on Computer supported cooperative work* (2002). 31–40.
18. Jerry Li, Mia Manavalan, Sarah D'Angelo, and Darren Gergle. Designing Shared Gaze Awareness for Remote Collaboration. In *Proceedings of the 19th ACM Conference on Computer Supported Cooperative Work and Social Computing Companion* (2016). ACM, 325–328.
19. 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 (2002), 257–278.
20. Joshua Newn, Eduardo Velloso, Fraser Allison, Yomna Abdelrahman, and Frank Vetere. 2017. Evaluating Real-Time Gaze Representations to Infer Intentions in Competitive Turn-Based Strategy Games. In *Proceedings of the Annual Symposium on Computer-Human Interaction in Play*. ACM, 541–552.
21. Diederick C Niehorster, Tim HW Cornelissen, Kenneth Holmqvist, Ignace TC Hooge, and Roy S Hessels. 2017. What to expect from your remote eye-tracker when participants are unrestrained. *Behavior Research Methods* (2017), 1–15.
22. Pernilla Qvarfordt, David Beymer, and Shumin Zhai. 2005. Realtourist–a study of augmenting human-human and human-computer dialogue with eye-gaze overlay. *Human-Computer Interaction-INTERACT 2005* (2005), 767–780.
23. Daniel C Richardson and Rick Dale. 2005. Looking to understand: The coupling between speakers' and listeners' eye movements and its relationship to discourse comprehension. *Cognitive science* 29, 6 (2005), 1045–1060.

24. Daniel C Richardson, Rick Dale, and Natasha Z Kirkham. 2007. The art of conversation is coordination common ground and the coupling of eye movements during dialogue. *Psychological science* 18, 5 (2007), 407–413.

25. John Sall, Ann Lehman, Mia L Stephens, and Lee Creighton. 2012. *JMP start statistics: a guide to statistics and data analysis using JMP*. Sas Institute.

26. Bertrand Schneider and Roy Pea. 2013. Real-time mutual gaze perception enhances collaborative learning and collaboration quality. *International Journal of Computer-supported collaborative learning* 8, 4 (2013), 375–397.

27. Kshitij Sharma, Sarah D'Angelo, Darren Gergle, and Pierre Dillenbourg. Visual Augmentation of Deictic Gestures in MOOC videos. In *12th International Conference of the Learning Sciences* (2016) *(ICLS'16)*. ACM.

28. Srinivas Sridharan, Reynold Bailey, Ann McNamara, and Cindy Grimm. 2012. Subtle gaze manipulation for improved mammography training. In *Proceedings of the Symposium on Eye Tracking Research and Applications*. ACM, 75–82.

29. Randy Stein and Susan E Brennan. Another person's eye gaze as a cue in solving programming problems. In *Proceedings of the 6th international conference on Multimodal interfaces* (2004). ACM, 9–15.

30. Stephen A Sweet and Karen Grace-Martin. 1999. *Data analysis with SPSS*. Vol. 1. Allyn & Bacon Boston, MA.

31. Yanxia Zhang, Ken Pfeuffer, Ming Ki Chong, Jason Alexander, Andreas Bulling, and Hans Gellersen. 2017. Look together: using gaze for assisting co-located collaborative search. *Personal and Ubiquitous Computing* 21, 1 (2017), 173–186.