

Look together: using gaze for assisting co-located collaborative search

Yanxia Zhang¹ · Ken Pfeuffer² · Ming Ki Chong³ · Jason Alexander² · Andreas Bulling⁴ · Hans Gellersen²

Received: 23 May 2016 / Accepted: 31 August 2016
© The Author(s) 2016. This article is published with open access at Springerlink.com

Abstract Gaze information provides indication of users focus which complements remote collaboration tasks, as distant users can see their partner's focus. In this paper, we apply gaze for co-located collaboration, where users' gaze locations are presented on the same display, to help collaboration between partners. We integrated various types of gaze indicators on the user interface of a collaborative search system, and we conducted two user studies to understand how gaze enhances coordination and communication between co-located users. Our results show that gaze indeed enhances co-located collaboration, but with a trade-off between visibility of gaze indicators and user distraction. Users acknowledged that seeing gaze indicators eases communication, because it let them be aware of their partner's interests and attention. However, users can be reluctant to share their gaze information due to trust and privacy, as gaze potentially divulges their interests.

Keywords Eye tracking · Gaze awareness · Gaze interaction · Multi-user · Collaborative task · Large pervasive display

1 Introduction

There are increasing numbers of high-density information large displays installed in public and work places. Recent research has shown that eye tracking can be employed in these pervasive displays [30–34]. These displays afford group activities, because a large display itself can act as a shared source of information used by multiple persons [23]. In a meeting, for example, a team of geologists can gather around a large map on a shared display to plan an upcoming trip. It is foreseeable that pervasive displays can track multiple users' gaze and enhance group interaction [19, 33].

Mutual gaze awareness is important in communication and collaboration in group activities. For example, in “backing away” scenarios when two users sit or stand at a distance from large displays to view the entire display and look for information together (see Fig. 1), gaze cues (e.g. eye contact and joint attention) provide rich context information that other body cues cannot reveal. To understand how gaze can enhance collaborative activities on a large shared display, we propose to provide visual representations of mutual gaze awareness into the design of a shared display interface.

Prior works proposed different ways to convey gaze cues visually. These include the use of video images of the partner's face and head [24, 27], gaze cursors [2], shared visual space (e.g. focused objects) [4], and scan paths overlaid on a screen [21]. These designs provide different gaze cues and are mostly targeted at remote settings. However, it is not clear what gaze cues are useful for in co-located collaboration. In addition, integrating gaze as visual representations on a shared user interface could potentially clutter the interface and interfere with group activities. This essentially raises another open question of how to present gaze cues effectively to benefit collaboration.

✉ Yanxia Zhang
Yanxia.Zhang@tudelft.nl

¹ Delft University of Technology, Delft, The Netherlands

² Lancaster University, Lancaster, UK

³ Hong Kong, China

⁴ Max Planck Institute for Informatics, Saarbrücken, Germany



Fig. 1 Gaze-assisted co-located collaborative search (arrows indicate gaze directions)

To address the above research questions, this paper presents an exploratory study to understand how gaze cues can enhance collaboration between two users in front of a large shared display. This paper first presents an implementation of our system that supports gaze visualisation of two users and the design of four gaze representations. We then present two empirical studies. In the first study, we examine how different gaze representations affect user performance and people's preferences in an abstract collaborative visual search task, where participants search for a specific object on a display with high-density information. The results show that people prefer a subtle and less explicit gaze representation to reduce distractions, but there is a trade-off between visibility and distractions. We further improve our gaze representation design based on findings from the first study and integrate it into a tourist map application (see Fig. 1). In the second study, we aim to understand the usage of gaze representation and subjective experience of the gaze-enhanced map application. We learn that gaze indicators can ease communication. However, some people are reluctant to share their gaze due to privacy concerns.

2 Related work

2.1 Gaze for multi-user interfaces

2.1.1 Eye contact for video conference

Gaze has been shown as an important cue for face-to-face communication [3, 6]. One of the major challenges in remote communication systems is to enable gaze awareness, because gaze cues can get easily lost in video conferences when users move freely in spaces. A plethora of research in HCI has investigated how gaze cues, mainly eye contact and mutual gaze, affect communication in video conferencing systems [24] and in immersive virtual environments [26]. One example of such systems, the GAZE

Groupware, conveys gaze in multiparty communication and cooperative work, such as in meetings [27]. Their work suggest that eye contact and gaze cues can help regulate conversation flow, provide feedback for understanding, and improve deixis in remote video conferences systems.

2.1.2 Gaze for remote collaboration

In collaborative work systems, the use of gaze has been investigated in remote setups. Similar to using gaze in remote communication systems, the *Clearboard* system enables gaze awareness between remote collaborators by using the metaphor of a transparent glass window [13]. Users are virtually located opposite each other to work on a shared board and can look through the transparent board to see what their partner is looking at. Although mutual gaze and the perception of eye contact can enhance the perception of co-presence, it seems to be far less important than the view of a group's shared work space on collaborative activities [7].

Some studies investigated the role of shared gaze in collaborative systems. The motivation comes from allowing remote collaborators to share their gaze over each other's screen space (i.e. seeing a collaborator's visual focus of attention). Previous research has pointed out that gaze plays a role as a "conversational resource" during spatial reference [14]. Gaze has been proposed to assist verbal collaboration in remote setups, due to the verbal communication problems like misunderstandings and noise. In a tourist planning application, Qvarfordt and Zhai applied gaze in a dialogue system [22]. They discovered that a remote assistant that is following remote users' gaze patterns while conversing with them can detect the users' interest. In a remote collaborative visual search task, Brennan et al. [2] demonstrated that sharing gaze is more efficient than speech for the rapid communication of spatial information. Similar results were found in [17] where shared gaze was shown to be more efficient than speech during collaborative tasks that require rapid communication of spatial information. Shared gaze has also been found useful to detect misunderstanding to overcome the lack of deixis at a distance [4].

2.2 Conveying gaze cues in collaboration

Based on prior findings in observation studies, we learn that multiple gaze cues can benefit collaboration on a large shared display.

Gaze has been considered as a valuable communication resource [14]. It naturally provides moment-by-moment information about a collaborator's focus, which can facilitate the interpretation of the partner's utterance because they can see the object that their partners are attending to. Seeing where the speaker is looking at has been found to

Table 1 Shared gaze in collaboration

Task types	Setup	Role	Mechanisms
Video conference	Remote [24, 27]	Regulate conversation	Faces and head
Problem solving	Remote [7, 25]	Understanding comprehension	Gaze cursor; scan path
Referential instruction	Remote [4]; Co-located [16]	Joint attention	Gaze cursor; Visual space
Visual search	Remote [2]	Spatial reference	Gaze cursor
Our work	Co-located	Communication; coordination	Four gaze representations

make disambiguation of their referring expressions early [25]. In particular, collaborations on a large display often involve members frequently referring to a specific piece of information on the shared display that is related to their discussion. The action of identifying on-screen objects is often carried out verbally, but when information is in high density, unstructured, and cannot be described using simple phrases, people may resort to body languages, such as pointing. Gaze can be a natural source of input information that benefits collaboration.

Another aspect in our face-to-face communication that gaze enables is to establish joint attention.¹ Achieving joint attention is critical for successful collaborative activities where groups reach a common ground in decision-making [5, 29]. As users gather around a large shared display, the eye contact and gaze cues can easily get lost due to different body orientation and focus changes between individual and group tasks [23, 28]; for example, when people stand or sit side by side in front of the display. This can make the process of establishing a joint attention challenging (see Fig. 1). Similar issue has been reported in previous study on collaborative data analysis on shared displays [12]. Their results revealed that participants commonly overlaid their mouse cursors to the joint focus area to show joint attention on a specific information item under discussion. Group members in their study further requested additional visual aids for drawing attention to mouse cursors.

Additionally, gaze can provide information that other body cues cannot reveal, such as ongoing cognitive activities (e.g. scanning, interests towards an object, and comparisons of different objects) [25]. These can potentially improve collaboration, as observing another person's gaze patterns might reveal the task status of the partner and gain information about other's intention.

2.3 Mechanisms for shared gaze

Prior research has proposed various ways of conveying gaze cues (see Table 1 for a classification of existing work). For example, video-mediated communication

systems show video images of the user's face to compensate for eye contact [24, 27]. Another common approach is to present users' gaze (i.e. shared gaze) as a cursor or focused object in the shared visual space, which helps them to be aware of their partner's focus [4, 17, 25]. Maurer et al. [16] proposed the use of co-driver's gaze cursor as a possible way of sharing information and fostering collaboration between driver and co-driver. Dynamic eye movements (e.g. scan paths) have also been found to enhance sharing of mental states [7, 21]. Enhancing gaze awareness in collaborative activities has been mostly investigated in remote settings (see Sect. 2.1).

The benefits of shared gaze in remote collaboration motivate our research. While previous works focused on remote settings, we further extend this notion in co-located collaboration on a large screen (see Table 1). Based on existing designs for shared gaze, we investigate how to provide gaze cues (e.g. direct visual attention and real-time eye movements) effectively and what effects they have on the collaboration.

3 System design and implementation

We implement our system using C# in Windows 8. Figure 2 illustrates the architecture of our system. We connect two Tobii EyeX/Rex eye trackers to a laptop (2.7 GHz, 16 GB RAM) that runs the system application, and the laptop is connected to an external large display (120 cm × 70 cm, 1080p resolution) for output. The eye trackers detect users' gaze at a minimum frequency of 30 Hz (i.e. every 33 ms).

When the eye trackers receive gaze data (Fig. 2), the system processes it in the following four stages:

Stage 1 Tobii SDK We use the Tobii Gaze SDK to extract raw gaze data from the eye trackers. The SDK provides gaze points (x, y coordinates with reference to the display), eye positions, head positions, and presence data. The data are then sent to the next stage to determine the users' fixation points. For each eye tracker, the system runs a dedicated process to receive gaze data. The gaze data values are sent via the signalR packages to the main Windows 8 Store App "controller" which is used to calculate the smoothed gaze data.

¹ *Joint attention* is when participants are mutually oriented to a common part of their shared visible environment and are aware that their conversational partners are also looking at it [29].

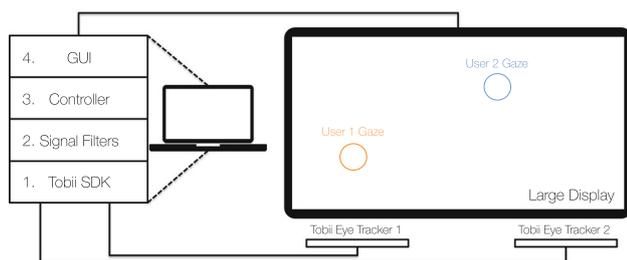


Fig. 2 The application receives gaze data from two eye tracking devices. Upon receiving the gaze data, the application first preprocesses the data and then informs the controller to update the positions of users' gaze visualisation on the user interface

Stage 2 Signal Filters Human eyes jitter during fixations because our eyes naturally make small involuntary movements (e.g. micro-saccades). Hence, raw gaze data are inherently noisy. To smoothen raw gaze data, we filter out saccade movements by calculating the real-time distance between gaze points. First, we compute the x - and y -axis displacements between current and previous detected gaze positions. Any gaze displacement (i.e. eye movement) that is above the distance threshold of 120 pixels is classified as a saccade, and otherwise is classified as a continuous fixation.

To further stabilise the fixation data, we use a weighted average to smooth the gaze data. Similar to [15], we calculate a fixation point in a time window of i frames (i.e. equivalent to approximately 500ms of gaze data) by using the following equations:

$$x_t = \frac{i * x_{t-1} + (i - 1) * x_{t-2} + \dots + 2 * x_{t-(i-1)} + x_{t-i}}{i + (i - 1) + (i - 2) + \dots + 3 + 2 + 1} \quad (1)$$

$$y_t = \frac{i * y_{t-1} + (i - 1) * y_{t-2} + \dots + 2 * y_{t-(i-1)} + y_{t-i}}{i + (i - 1) + (i - 2) + \dots + 3 + 2 + 1}, \quad (2)$$

where i represents the window size ($i = 15$ in our case).

The current fixation point is sent to the controller as an event to update the previous fixation.

Stage 3 Controller When the controller component receives a fixation point, it updates the position of the corresponding gaze object (e.g. a cursor). In other words, if new gaze data are received from eye tracker 1, then the gaze object for tracker 1 is updated. This changes the x and y coordinates of the gaze object on the Cartesian plane of the display.

Stage 4 GUI Lastly, the application informs the system to render any updated gaze-controlled objects on the display at 10 Hz. We do this to maintain a smooth refresh rate due to irregularity from the fixation data.

During our pilot trials, we test several configurations of thresholds and window frames. Although the current implementation has a delay of one frame (i.e. 33 ms), it enables a more stable focus point representation and also allows fast shifts between fixations.

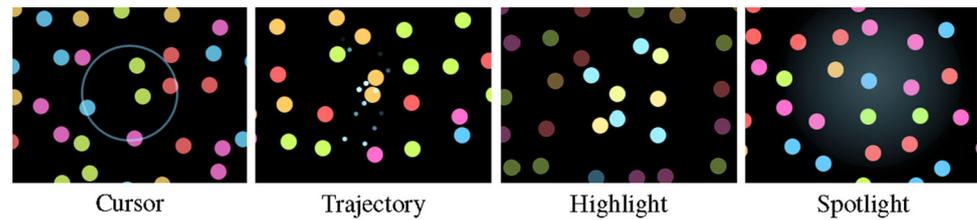
3.1 Gaze representation design

In this work, we present four types of gaze representations that aim to support users in co-located collaborative tasks based on existing designs summarised in Table 1 (Fig. 3):

- **Cursor Gaze** is displayed as a coloured circular ring with a radius of 60 pixels. This type of gaze representation is similar to having an onscreen cursor following a user's gaze. This is consistent with the gaze cursor in Table 1.
- **Trajectory Gaze** data within the last 3 s are plotted as a trajectory. Each sample is displayed as a small circle, and its opacity decreases with time. Hence, the most recent gaze data have the highest opacity. Trajectory is a representation of the scan path in Table 1.
- **Highlight** Displayed objects within a 60-pixel radius from the gaze point are highlighted by increased brightness. Any objects that are nearby the user's gaze will be automatically made more visible or selected. This is similar to the visual space on focused objects in Table 1
- **Spotlight** This simulates a torch shining effect (shown as a bright Gaussian-blurred disc) that follows the user's gaze location. The resolution is full in the central fovea within 2° of visual angle and falls gradually towards 3 degrees beyond the periphery. This simulates human visual perception. Its resolution is much higher at the fovea focus than the periphery [18], and hence, Spotlight's opacity gradually fades from fovea to periphery. This is similar to the visual space on focused objects in Table 1.

4 Study 1: Effects of gaze representation

In this study, we aim to evaluate how people perceive the usefulness of the four gaze representations as communication and coordination cues on a shared display. The goal is to investigate how different representations of gaze help collaboration. We selected a visual search task adapted from Brennan et al. [2]. Participants collaboratively search for an oval object amongst a large set of non-overlapping circular objects. They are required to make a joint decision to confirm or reject whether the oval object exists. The task has similar elements as real-world collaborative visual search tasks, where people would need to look for information together in front of a high-density display, such as

Fig. 3 Four types of gaze representations

locating a specific building on a campus map, or finding a particular product in a shopping catalogue.

In our study, we aim to understand the following research questions:

- Can gaze representations improve users' performance in collaborative search tasks?
- Can gaze representations influence people's perception of communication and coordination in collaborative tasks?
- Do people feel distracted or attentive when seeing different gaze representation designs? How do they influence collaboration?

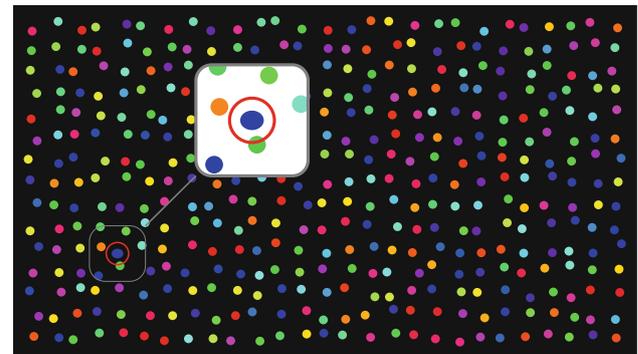
We hypothesise that providing gaze information of collaborators can help them to become more aware of each other's attention, and thus better facilitate their communication to reach a common ground. We further hypothesise that gaze history in temporal space (like gaze trajectory) would provide collaborators with revealing additional information of their partner's attention and search strategy, and thus better coordinate their search actions.

4.1 Participants and setup

We recruited 16 participants (13 males and 3 females, with a mean age of 27.9 years SD 4.7 years), as 8 pairs to take part in the study. We used a 55-in display (120 cm × 70 cm, 1080p resolution), with the bottom bezel positioned at a height of 115 cm above ground. Each pair of participants stood side by side and at a distance of 2 m in front of the display, with a view angle of 46.4° horizontally and 28.1° vertically. Two eye trackers were placed at a distance of 140 cm in front of the display, each tracking one user's eyes. One eye tracker was placed at 30 cm to the left of the screen's centre; the other one was placed at 30 cm to the right. The eye trackers were aligned at a height of 5 cm above the bottom of the screen. We conducted a pilot study to fine-tune setup parameters, such as the sizes of gaze representation. We found that a 60-pixel radius (3 degrees of visual angle) is the optimal size.

4.2 Task and procedure

The participants' task is to make a joint decision of whether they find a coloured oval target (0.8° in height and 0.95° in

**Fig. 4** Study 1: visual search stimulus

width) amongst 364 non-overlapping coloured circles (0.8° visual angle). Each task consists of one of two conditions: *target-present* or *target-absent*. The target-present condition consists of one oval target, placed in a random non-overlapping location amongst other circular dots. In the target-absent condition, all dots are circles (Fig. 4).

We adopt a within-subjects design for five conditions: *without gaze*, *gaze cursor*, *gaze trajectory*, *objects highlighting*, and *spotlight* (see Fig. 3). In the without gaze condition, the display provides no gaze visualisation. In the other four conditions, both participants see where they are looking at in real-time on the screen, and the gaze visualisation is colour-coded for the respective users (orange, blue). The order of the five conditions was counterbalanced. Each study session consisted of 60 trials (hence 12 trials per gaze visualisation condition), and half of the trials were target-present.

Prior to the study, the eye trackers were calibrated individually to each participant. Participants were allowed sufficient time to practise. A 3-min break was given after completion of each condition (i.e. 12 trials).

The participants were asked to complete the task as fast and accurately as possible. They were allowed to converse freely with their partner, without restrictions on strategy or communication. After the first participant responded, they received feedback about the correctness. Each session lasted approximately 60 min.

4.3 Data collection

We collected quantitative and qualitative data. During the study sessions, the system logged the participants'

completion time of each trial and the number of errors made for each condition. After completing each condition, the participants answered questionnaires which made up of 7-point Likert scale questions and open-ended questions for their subjective experience. We balanced the Likert scale questions with both positive and negative questions.

The questionnaire consists of multiple parts. The first part focuses on how people perceived the quality of collaboration and the mental and physical effort required to use gaze indicators for collaboration; for example, how gaze representation helps them to make joint decisions, as well as assists communication and coordination between partners. The second part focuses on the effectiveness of gaze feedback, and we ask questions that are related to distractions, usefulness, and whether and how gaze indicators hinder collaboration.

The questionnaire also asks participants about the strategies that they adopt for collaborating with their partner to complete the task, such as the types of difficulties that they encountered, what types of information that the participants gain from seeing the partner's gaze indicators, and how they feel about the value of seeing the gaze indicators.

Lastly, the experimenter conducted a short interview with the participants (as a pair together) for feedback and suggestions for improvement about the effects of different gaze representations.

4.4 Results

4.4.1 Group performance

We measured the overall search time and accuracy for each visualisation condition. Figure 5 illustrates the average search times for the target-present and target-absent trials. The results of average search accuracy across the different gaze representations are presented in Table 2. The average search accuracies for different conditions are similar.

A repeated measure ANOVA analysis showed a significance for completion time across the five conditions in the target-absent ($F(4,28) = 2.728, p < 0.05$) trials and in the target-present trials ($F(4,28) = 2.762, p < 0.05$). However, pairwise comparisons showed no pairs with a significant difference in the target-absent trials. Spotlight achieved the shortest completion time in target-present conditions. A significant result ($p < 0.05$) was obtained in target-present trials, with the Spotlight ($M = 14.6$ s) being faster than the None ($M = 21.7$ s) condition. Our data showed that gaze information can improve the speed of the collaboration task; however, the way of presenting gaze feedback can influence people's performance in speed.

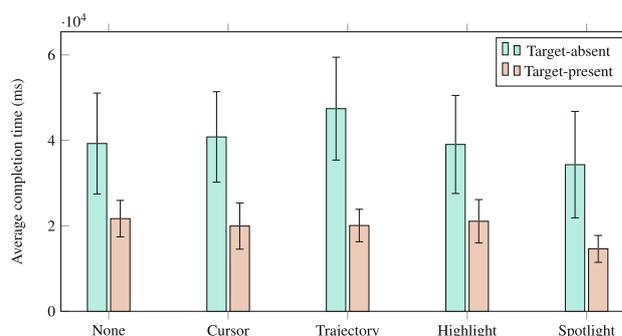


Fig. 5 Average of the overall search time. Error bars represent the 95 % confidence interval of the mean

Table 2 Average search accuracy

	None	Cursor	Trajectory	Highlight	Spotlight
Mean (%)	81.7	83.3	80.8	81.7	80.8
Std (%)	17.7	17.7	24.2	14.2	19.2

4.4.2 Gaze role: feedback and observations

Gaze for communicating spatial information Half of the participants (8/16) mentioned that seeing the gaze indicator was helpful and it became “easier to explain to each other where the target was”. Gaze was more convenient than speech to describe a target position (such as pointing out a particular display region and colour). After getting used to having gaze visualisation, some participants commented that “it was strange not to have any indicator of my partner's gaze” in the *None* condition. Subjective feedback also revealed that users found the gaze indicator useful to indicate the location of a target. Without gaze information, people needed to speak more to explain the location of a target, and they found it *easier* to communicate with gaze indicators. For some participants, gaze information was particularly useful when they needed to confirm or come to an agreement with their partner.

Gaze for coordination The participants had diverse ways for coordinating the search strategies. When users searched together, they first started with establishing rules by verbal communication. For example, the majority of our participants started with splitting the screen in two regions, like “I start right, you start left” or “I [go] left to right and my partner [goes] top to bottom”.

An interesting observation we noticed is that, when gaze information was shown, people tended to avoid looking at the same region together at the same time, and this was usually done without explicit verbal communication. For example, if a user saw that his partner was searching the top-right region, the user would choose another region to search. One of our participants explained, “the gaze

indicator showed where my partner was looking, so I could look at other parts of the display”. This minimised the chance of both users doing the same thing simultaneously, as gaze indicators made them aware of their partner’s progress. Other times, users synchronised their actions with the partner, for example, “First we focused on different sides (left and right) next we scanned the middle part together”. Thus, they first split the workload and then combined.

The questionnaire data also reflected that the users were monitoring their partner’s focus and attended areas through the partner’s gaze indicators (e.g. by their peripheral vision). The intention of keeping themselves aware of the partner’s gaze was mainly due to the participant adapting their search strategies to cooperate with the partner. Some participants mentioned that they defined a strategy beforehand, hence to gain progress by checking where their partner was looking. For instance, in between if they found their partner’s gaze indicators appearing in their half and they would wonder if the partner was properly searching his half and if he “should check his [the partner] half too”.

Gaze for attention guide Users occasionally lost track of their searching location due to distraction or tiredness. In the gaze trajectory condition, several participants expressed how they used their gaze indicators as a guide for finding where they were scanning. Our participants commented, “sometimes I got confused about where I was, but because of this indicator, I can quickly continue from where I [got/was] lost”. The tail of gaze trajectory provided implicit information of the user’s scanning process, so when the user was distracted they could quickly refer back to the trajectory tail to continue.

4.4.3 Effects of the gaze feedback

The majority of participants did not consider that the task was difficult to complete collaboratively with their partner in the *None*, *Highlight*, and *Spotlight* conditions (see Fig. 6). A third of the participants agreed that the

Trajectory condition made the task more difficult than the other conditions. Similarly, the *Trajectory* condition was consistently rated higher for physical demand than the *None* condition. Our questionnaire data suggest that the physical demand was mainly induced by eye fatigue. However, a Friedman test on users’ responses (with regard to difficulty to complete the task, mental demand and difficulty in communicating and coordination on all conditions) did not reveal a significant difference (see Fig. 6).

When we asked the participants about problems and difficulties that they encountered, we learned that the major difficulty was from the presence of the gaze indicator during the normal viewing process, which often distracted them from visually searching. When looking at the user feedback about the effects of different gaze feedback, there is no significant result found in any particular representation winning over the other technique (compared using the Friedman test; see Fig. 7). Participants agreed that seeing the gaze indicators was distracting in *Cursor*, *Trajectory* conditions, while the object *Highlight* and *Spotlight* conditions were less distracting.

In the *Cursor* condition, eight participants mentioned that they felt the gaze cursor was distracting although they found it easy to make an agreement in this condition. One problem encountered by many participants was the occlusion by the gaze cursor which made it hard to judge the oval target shape. Other problems include that the cursor was “inaccurate” and “moving too much” which was caused by instability of human fixation, and the cursor “size [was] too big”.

In the *Trajectory* condition, five participants found this representation very distracting which made the search task difficult. They commented that “the movement [of the trajectory] is very distracting”, in particular, when two tails (from two users) crossed each other. The side effect was that the participants could not accurately and precisely infer where the other was looking at, rather being unintentionally chasing the other’s gaze from time to time. In some cases, the participants even tried to scan faster than

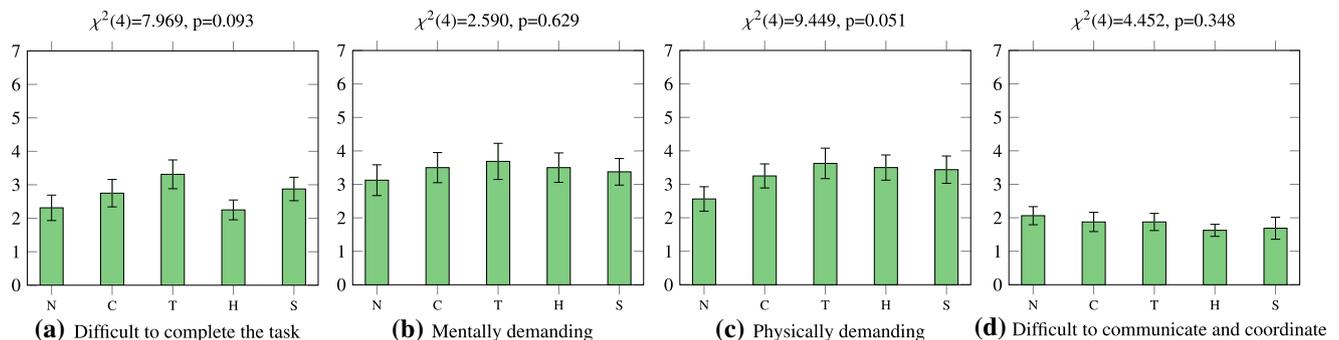


Fig. 6 Subjective feedback on collaboration experience to complete the search task (1-Strongly disagree to 7-Strongly agree). The error bars in all figures stand for the standard error of the mean. N (None), C (Cursor), T (Trajectory), H (Highlight), S (Spotlight)

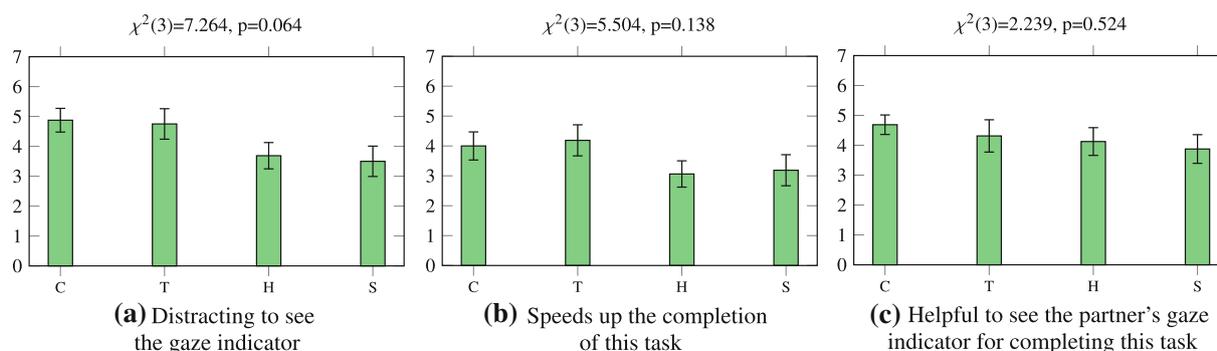


Fig. 7 Subjective feedback on effects of the gaze feedback (1-Strongly disagree to 7-Strongly agree). The *error bars* in all figures stand for the standard error of the mean. N (None), C (Cursor), T (Trajectory), H (Highlight), S (Spotlight)

the cursor to evade the problem. It seems that the advantage of using gaze for spatial referencing decreased in the *Trajectory* condition, as this type did not provide precise representation of current focus location. Hence, participants felt that it only indicated a rough region and they still needed to perform a further search to locate the target. On the other hand, three participants found this type of gaze indicator helpful as it revealed the partner's search speed, so that they could adjust to cooperate.

In the *Highlight* and *Spotlight* conditions, the majority of the participants felt the indicator was less distracting, e.g. *very subtle and not distracting*. They felt that they could focus on searching and still know what their partner was looking at. The only problem encountered for the *Highlight* feedback was the glimmer effect (mentioned by two participants). In the *Spotlight* condition, two participants mentioned that they felt the indicator was like “a proper element that was on top” which sometimes caused them to focus on the gaze feedback rather than the stimulus. As these two types of gaze feedback were more subtle with less visibility, the effects of assisting target referencing were less prominent (see Fig. 7b, c).

4.5 Lessons learned

When is gaze useful?: From this study, we learned that that gaze information can be useful in the collaborative search task in a co-located setup on a shared screen, e.g. for referring a remote target, being aware of a partner's focus and guiding their own attention. The gaze information would benefit in particular when people need to cooperate and coordinate with their partner. Although participants mentioned that it was useful and interesting to keep an eye on where their partner was looking, gaze was found to be less useful during the normal searching and viewing process. It is still unclear whether users would need the gaze information all the time during their collaboration or whether it would distract them more from their individual goal.

Avoid gaze trajectory: Our results suggest that the *Trajectory* feedback should be avoided in scenarios where frequent target referencing is required. The main difficulty came from the irregularity of the generated gaze trajectory patterns. The characteristics of eye movements (e.g. saccades) were different from continuous pointer movement such as mouse. Thus, the created trajectories varied in shapes and lengths depending on the amplitude and speed of the eye movements. This non-uniform representation confused users and was less useful in both cases for assisting spatial reference and communicating attention.

Subtle gaze feedback (visibility vs. distraction): One of the biggest challenges we realised is the conflict between visibility and distraction of the gaze indicators. High visibility gaze indicators (e.g. cursor and trajectory) provided fast and accurate target reference, however, caused more distraction. Users preferred subtle representation of gaze feedback in the object highlighting and spotlighting representation. Representing gaze as an object (e.g. a cursor) can distract users. However, when the visibility decreases, the gaze indicator loses its power for spatial referencing and maintaining focus and attention awareness during the collaboration.

5 Study 2: Tourist map application

Our second study investigates people's qualitative experience in a more realistic setup. We built a tourist map application like those in information centres, train stations, or museums (Fig. 8). Two users communicate and find a hotel on the map that they both agree and approve to.

Our application integrates two gaze visualisations. From the previous study, we learned that people prefer gaze visualisations that are subtle and less conspicuous, e.g. the highlight and the spotlight gaze representations. We combine the two types of visualisations into a single gaze indicator as illustrated in Fig. 8b.

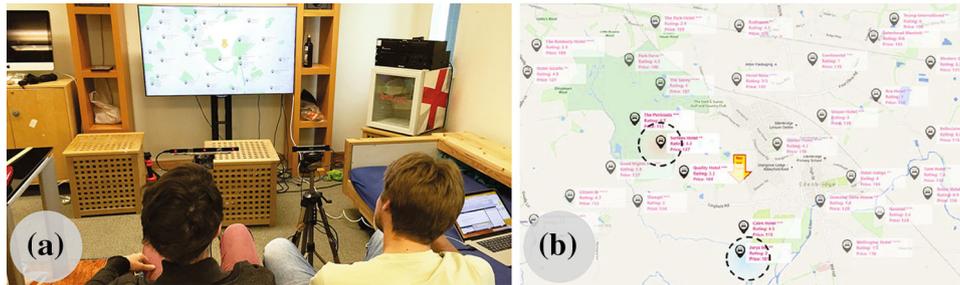


Fig. 8 Setup: **a** A pair of participants sat in front of a large screen, with an eye tracker facing each person to capture their eye movement. **b** The application interface showing the gaze indicators of two users (the *dashed circles* are not part of the interface; only added for visibility)

We also added a foot control that enables users to switch the gaze visualisation on or off. Our first study showed that gaze indicators can be distracting from time to time, and we thought to provide the user with more control of their gaze visuals. We chose a foot control so that the user's hands are kept free, enabling natural use of hands for body language during discussion, and to potentially hold on to private items during the activities (in contrast to hand-based control such as mouse/keyboard).

5.1 Study design

We recruited 20 participants (10 pairs, 16 males, 4 females, age from 21 to 43, $M = 29.7$ $SD = 5.8$) from our research department. The setup was similar to the first study, except this time the participants were seated instead of standing (Fig. 8a).

Prior to the study, we demonstrated our tourist map application to the participants and allowed them sufficient time to calibrate the eye trackers, to experience the interaction, and to get comfortable with the system. The system presents a map with 30 hotels (chosen randomly from a pool of 75 hotels) scattered across the screen (Fig. 8b). Each hotel is attached with its name, hotel quality rating (i.e. number of stars), price, location, and average customers rating (on a scale out of five).

During the study, we explained to the participants that they should assume that they are tourists who are travelling together and looking for a hotel. The participants were free to discuss with each other. Their task was open-ended, and the only requirement was that they must come to an agreement of selecting a hotel. To stimulate discussion, each participant was advised to look for hotels that satisfied specific conditions. For example, one participant would look for nearby hotels that are close to where they are (indicated by a “You Are Here” maker), while the other participant would seek for hotels with a good reputation (e.g. user rating).

On average, a study session lasted for approximately 30 min, and every session consisted of eight trials. For each trial, a random map was loaded with new hotel information. After four trials, the default settings inverted. After

completing the eight trials, the participants filled in an exit questionnaire with their subjective feedback. Half of the participants started with gaze indicators being switched on by default, and the other half with gaze indicators switched off initially. This helps us to learn when users would invoke the gaze indicator and in what situations they would want to make the indicators hidden or visible.

5.2 Data collection

We collected system logs and qualitative feedback through a two-part questionnaire. The first part focused on the participants' collaboration experience. We elicited their feedback by asking questions about how the gaze indicators assisted them to collaborate with their partner. In conjunction, we used an adaptation of the desirability toolkit [1]; we provided the participants with a list of adjectives and asked them to select five or more that most closely matched their personal reactions to the system. The method of selecting adjectives is ideal to elicit a participant's reactions and attitudes, as it provides a quick high-level indication of their reactions. The selection of words then acts as a basis for further explanation and elaboration about why they chose those words.

The second part of the questionnaire focused on how the participants controlled the visibility of their gaze indicators. We asked questions on what caused the user to turn their gaze indicators on and off, as well as what caused them to avoid toggling the gaze indicator. This can help us to find out when the participants perceive gaze indicators as useful or counter-productive. Lastly, we asked the participants to identify any problems that they encountered during the study, the types of applications that they thought gaze indicators would be useful, as well as suggestions for future improvement.

5.3 Results

Our participants were positive on the use of gaze for collaboration. Most of them state that it was “convenient” to see their partner's gaze location, because it made them

aware of which location that their partner was referring to during discussion, and gaze also makes pointing at a map location simpler. Gaze enabled the participants to spend more effort on discussion instead of thinking of words to describe a specific location, as the users can simply point by staring. One participant mentioned that he preferred to describe a map location by referencing nearby landmarks, but acknowledged that gaze indicators are useful in “quiet” locations that had no nearby reference landmarks. The participants also mentioned that having the gaze indicators in different colours made them easily distinguishable and reduced confusion. However, a participant stated that any patches of background that had similar colour to the gaze indicator colour could make spotting the indicator difficult.

Several issues were reported, e.g. inaccuracy, which was caused by eye tracking detection errors. Some participants experienced a small distance offset between their focus and their gaze indicators for which they compensated by slightly looking off target. A few users also found their partner’s rapid gaze indicators to be distracting and needed to be conscious not to follow them. They suggested that gaze indicators should be less conspicuous and only be revealed on demand. While some people preferred less apparent gaze indicators, some actually preferred them to be larger and more visible. They explained that increasing visibility would help to explicitly catch other’s attention.

5.3.1 Reactions to gaze indicators

The participants agreed that having gaze indicators for collaboration was interesting (20/20) and the majority considered it pleasant (10/20) because the interface was “easy to learn” and provided a “straightforward experience”. The participants also stated that the gaze indicators made the task more efficient (15/20) as it provided an “extra layer of information between [the partners] ... by just looking at [the target]”, and smooth (8/20) because the “[gaze indicator] followed the eyes ...and saved complicated location description”. Several people mentioned that the experience could be stressful (5/20) because of the distraction of the gaze indicator, so the users needed to “[focus] on the pointer all the time”. The experience could also be frustrating (3/20) due to inaccuracy which caused the interaction to be “chunky”, “jerky”, and “slow to get the pointer to the exact location”.

5.3.2 Gaze indicators for collaboration

The participants frequently described their experience of having gaze indicators for collaboration as helpful (14/20). The primary benefits pointed out by the participants are that using the system was time-saving (12/20) and it speeded up the interactions. At the same time, participants

also felt the collaboration experience to be fun (9/20) and entertaining (8/10). One participant even summarised his experience as “a tedious and potentially worrisome task made easy, pleasant, and efficient”.

The participants considered that the interface was simple (13/20) and intuitive (7/20). They acknowledged that gaze indicators can enhance communication, as the users are made aware of their partner’s interests. They also recognised that the gaze indicator reduces effort and shortens verbal description, since the gaze already acts as an immediate pointer.

At the same time, gaze indicators also helped the users to gain an idea on whether their partner was paying attention to what they were talking about. We observed an instance where one participant stepped on his partner’s foot control to turn on the partner’s gaze indicator, so he could know where the partner was looking. Several participants also felt that using the system was frustrating (3/20) and overwhelming (2/20). Sometimes it was because the participants needed a while to realise which gaze indicator belonged to whom. Other times it was caused by requiring attention to divert other’s focus to their gaze indicator while not following the partner’s gaze indicator.

5.3.3 On/off toggle behaviour

We observed two phases of collaboration. In the first phase, *scanning*, the participants individually looked for hotel options in parallel. Some participants considered that having the gaze indicators switched on during this phase could cause distractions. The second phase consisted of *discussion*. The participants often needed to refer to different hotel options on the screen and also to direct their partner’s attention to where they were looking. In the second phase, gaze indicators were frequently used, and the participants often switched the gaze indicator on to ensure that it was available. We also observed cases of frequent toggles of the gaze indicators when the participants wanted to refer to different on-screen targets during their discussion.

Three quarters (15/20) of the participants left their gaze indicators on and never switched them off. They explained that the gaze visualisation helped them to focus on picking a hotel option and also made it easier for their partner to see their preferences. Infrequently, five participants switched their gaze indicator off and explained that this was due to fatigue and distraction or they simply no longer wanted to search anymore. Inherently, switching the indicator off can be a social sign to inform the partner that they want to finish the task.

We also observed that some people switched their gaze indicator off for a brief moment and immediately turned it back on. This happened during their discussion of hotel

options, where the participants realised although distracting, the indicator was needed for a more efficient communication (like pointing at a hotel). There are also occasions that, when the gaze was off, people toggled their gaze indicator on for a brief moment and immediately toggled back. These instances happened when the participants wanted to use their indicators to quickly direct their partner's attention to what they were looking at when they want to pick another hotel option together. Several participants intentionally switched their indicator off because they were not comfortable and reluctant to let their partner see where they were looking.

6 Discussion

As collaborative activities often happen around a large shared display (e.g. surface hub, digital board), we believe that many collaborative applications can benefit from our studies. Our results show that having gaze with visual representations as implicit indicators of visual attention displayed on a shared display enables co-located partners to be aware of each other's focus and indeed helps them to communicate during collaborative tasks. We also show that different types of visualisation of gaze indicators can impact collaboration.

In this work we learned that:

- The subtlety of visual representation of gaze indicators influences the quality of collaboration. Highly visible visualisation can lead to distractions and hampers collaboration. Subtle and less explicit gaze representation is preferred.
- Displaying gaze indicators improves the efficiency of collaborative tasks, as users can refer to a specific on-screen location by looking. This eliminates the verbose process of describing the location verbally.
- Revealing gaze information enhances group synchrony and avoids duplication, as users are aware of collaborators' focus. A gaze indicator also helps to establish joint attention, which benefits collaborators' communication and understanding between partners.

6.1 Comparison with existing works

In conventional desktop settings, to convey users' focus of attention in shared workspaces, previous research proposed the use of visual representation of mouse movements (e.g. telepointers [8]) and integrating a variety of awareness widgets into the user interface. However, mouse cursors do not represent the users' focus, as cursors can be stationary while the users are paying attention to another location. In other words, cursors do not provide an accurate

representation of user attention. Visual awareness widgets (e.g. radar view), which are also determined from mouse cursor positions, require additional space of the shared workspace [11]. What we proposed in this work is to harness gaze as a natural information source of user attention to assist collaboration, which requires no extra user actions. In addition to presenting users' attention, gaze is also a natural pointer, so people can use it to provide spatial references and establish joint attention.

Similar to previous findings in workspace awareness research [9, 20], making actions more perceivable aids maintaining awareness. However, presenting more additional information can increase distractions. We encountered similar problems in our study. Although we found people in general prefer subtle gaze feedback (e.g. highlighting objects), in some cases people actually preferred obvious representations (e.g. spotlight). This happened because making gaze indicators obvious can be useful for spatial referencing and invoking the other's attention.

Our choice of task that is similar to Brennan et al. [2]. Brennan et al. focused on coordination aspects of gaze sharing, with respect to speech communication in a remote visual search task [2, 17]. Gaze was found to be superior to speech in terms of communicating spatial references. Interestingly, they found that using speech with shared gaze was substantially less efficient than using shared gaze alone due to the coordination cost of speech communication. On the contrary, with a different setup, in co-located settings, gaze enhances communication and coordination with body languages or voice cues. Also, we found that collaborators' gaze provides awareness information so that users would divide their tasks. What we often observed is that use of speech and the gaze indicator worked simultaneously to assist collaboration. Sometimes, speech was used to provide explicit instructions to coordinate action, while gaze was used as an implicit cue to decide the working area or to monitor the other's progress. Other times, gaze was used to initiate attention from the partner, whereas speech was used to confirm he is in the right place. However, the simultaneous use of the gaze indicator with hand gestures has been seen infrequently. This is probably because the gaze and hand gestures can similarly act as a pointer.

We further contribute the user experience aspects of sharing gaze in collaborative activities that have not been covered in previous research. Our results indicate that users had a positive experience with our shared gaze interface. The results are encouraging and our work opens further research opportunities for studying how gaze cues can be integrated into large displays to support more complex collaborative tasks. In future, we intend to study how gaze enhances other activities. For example, in a multi-device ecology, we often find many co-located collaboration

opportunities (e.g. cross-device interaction). We predict that gaze can show further benefits in scenarios when hands are occupied with manual input devices (e.g. mobile devices) and there require frequent changes of focus between group and individual devices/tasks.

6.2 Lessons learned and design considerations

Our proposed design is simple to implement and can be applied in many shared display applications. We encourage interface designers to consider our approach to use gaze for multi-user collaborative applications. In the following section, we provide lessons learned from this work and limitations of applying our approach.

6.2.1 Trust and privacy of shared gaze

In collaborative tasks, people often first agree upon a divide-and-conquer strategy, so that each person works on an individual region (e.g. one person focuses on the left, while the other focuses on the right). We observed that some people cross over and deviate to their partner's region for double checking. Having gaze indicators switched on can negatively impact partnership. Seeing a partner's gaze on a non-allocated region can be implied as a lack of trust or that the person is not following agreed instructions.

People naturally look at objects that they are interested in. By observing users' gaze indicators, it is possible to infer their interests. This poses a privacy concern, and users may not be willing to reveal their gaze focus, especially to strangers or to people whom they are not familiar with. In our second study, we provided a control feature for people to hide their gaze indicators. We observed that people would turn off the gaze indicator if they were uncomfortable about letting their partner know what they are looking at. Keeping gaze indicators on throughout the interaction may be acceptable when working with a trusted partner; however, the situation could differ if it is in a public environment. This inherently opens the question of *under what context and constraints are inappropriate to reveal gaze indicators?*

6.2.2 Augmented gaze representation

Integrate Semantic Information Similar to [20], the identity problems can cause distraction and confusion, especially with conspicuous gaze indicators that people often need to check which indicators belong to whom. This issue could be alleviated by adding identifiable denotation using strategies similar to telepointers [8], like attaching names, assigning different shapes, photos or arbitrary information to each user's gaze indicator.

Additional Visualisation Control In our design of our application, we only provide a function to toggle the visibility of the gaze indicator. Our observations helped us to realise that gaze indicators provide multiple benefits in assisting collaboration. Sometimes users prefer explicit and use the gaze indicator actively. But from time to time, people use it rather passively for monitoring the other's attention. It may be necessary to empower the users with some level of control over adjusting their gaze presentations. One solution could be, similar to the control of virtual embodiments in tabletop groupware systems [20], allowing users to actively adjust the opacity of visual representations.

6.2.3 Issues of eye tracking

Going Beyond a Pair In the setup of our study, we used an eye tracker for each user, because current commercial eye trackers can only support gaze detection of an individual user. This inherently constrains the number of simultaneous users. We envision that in the near future eye trackers can support simultaneous gaze tracking of multiple users. This essentially raises a new research question of *what happens if the interface presents many gaze indicators?* From the studies we learned that users get distracted easily from simply two gaze indicators. Increasing the number of indicators can intensify distractions. Although our users suggested that they prefer to have customised and distinguishable indicators to reduce confusion, finding the right balance between the number of simultaneous gaze indicators and the design of subtlety is an important aspect for future gaze-assisted co-located collaboration.

Stability of gaze representation Our experience informed us that eye movement patterns (using trajectories) are difficult to interpret in real-time. In addition, one of the biggest distractions, compared to visual representation used in other groupware work [8, 10, 11, 20], is actually from the jitteriness of the visual representation. In our work, we showed a simple threshold-filtering technique to remove saccades and to smoothen gaze raw data. We anticipate that more sophisticated fixation and saccade detection algorithms can improve the gaze stability.

7 Conclusion

This paper investigated the use of gaze for collaborative search applications. We presented two users' gaze locations (using four different representations) on the same display, to help collaboration between partners. Our results show that gaze can enhance co-located collaboration and help users' to coordinate their search strategies to minimise chances of doing the same work. However, there is a trade-

off between visibility of gaze indicators and user distraction. Users preferred subtle feedback such as using object highlighting and blurred gradient visual representations. Although gaze cursor and moving trajectory provided gaze information with high visibility, they seemed to be more distracting and less preferred by the users.

With a gaze representation design that combined both object highlighting and blurred gradient visual representations, users acknowledged that seeing gaze indicators eases communication, because it makes them aware of their partner's interests and attention. Users found gaze is helpful and time-saving when collaborating with partners. Users also perceive the use of gaze for communication easy and intuitive. We believe that the advantage of supporting gaze in co-located collaborative tasks can be further improved by appropriate design and considering how best to present gaze information to balance visibility and distraction.

Application designers should also take into account the issues of trust and privacy for gaze sharing. Besides interface aspects, users can be reluctant to share their gaze information due to privacy, as gaze behaviour is hard to fake and potentially divulges their interests.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Benedek J, Miner T (2002) Measuring desirability: new methods for evaluating desirability in a usability lab setting. In: Proceedings of UPA conference
- Brennan SE, Chen X, Dickinson CA, Neider MB, Zelinsky GJ (2008) Coordinating cognition: the costs and benefits of shared gaze during collaborative search. *Cognition* 106(3):1465–1477
- Broz F, Lehmann H, Nehaniv C, Dautenhahn K (2012) Mutual gaze, personality, and familiarity: dual eye-tracking during conversation. In: RO-MAN, 2012 IEEE (Sept), pp 858–864
- Cherubini M, Nüssli M-A, Dillenbourg P (2008) Deixis and gaze in collaborative work at a distance (over a shared map): a computational model to detect misunderstandings. In: Proceeding of ETRA 2008, ACM, pp 173–180
- Clark HH, Wilkes-Gibbs D (1986) Referring as a collaborative process. *Cognition* 22(1):1–39
- Cook M (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. *Am Sci* 65(3):328–333
- Gergle D, Kraut RE, Fussell SR (2004) Language efficiency and visual technology: minimizing collaborative effort with visual information. *J Lang Soc Psychol* 491–517
- Greenberg S, Gutwin C, Roseman M (1996) Semantic telepointers for groupware. IEEE Computer Society. In: Proceedings of OZCHI, pp 54–61
- Gutwin C, Greenberg S (1998) Design for individuals, design for groups: tradeoffs between power and workspace awareness. In: Proceedings of CSCW, ACM, pp 207–216
- Gutwin C, Greenberg S (2002) A descriptive framework of workspace awareness for real-time groupware. *Comput Support Coop Work* 11(3):411–446
- Gutwin C, Roseman M, Greenberg S (1996) A usability study of awareness widgets in a shared workspace groupware system. In: Proceedings of CSCW 1996, ACM, pp 258–267
- Isenberg P, Carpendale S, Bezerianos A, Henry N, Fekete J (2009) Coconutrix: collaborative retrofitting for information visualization. *IEEE Comput Graph Appl* 29(5):44–57
- Ishii H, Kobayashi M (1992) Clearboard: a seamless medium for shared drawing and conversation with eye contact. In: Proceedings of CHI, ACM, pp 525–532
- Kraut RE, Fussell SR, Siegel J (2003) Visual information as a conversational resource in collaborative physical tasks. *Hum Comput Interact* 18(1):13–49
- Kumar M, Klingner J, Puranik R, Winograd T, Paepcke n (2008) Improving the accuracy of gaze input for interaction. In: Proceedings of ETRA 2008, ETRA '08, ACM (New York, NY, USA), pp 65–68
- Maurer B, Trösterer S, Gärtner M, Wuchse M, Baumgartner A, Meschtscherjakov A, Wilfinger D, Tscheligi M (2014) Shared gaze in the car: towards a better driver-passenger collaboration. In: Adjunct proceedings of automotiveUI, ACM, pp 1–6
- Neider M, Chen X, Dickinson C, Brennan S, Zelinsky G (2010) Coordinating spatial referencing using shared gaze. *Psychon Bull Rev* 17(5):718–724
- Perry JS, Geisler WS (2002) Gaze-contingent real-time simulation of arbitrary visual fields. *SPIE Proceedings*. In: Human Vision and Electronic Imaging, pp 57–69
- Pfeuffer K, Zhang Y, Gellersen H (2015) A collaborative gaze aware information display. In: Adjunct Proceedings of UbiComp, UbiComp/ISWC'15 Adjunct, ACM, New York, pp 389–391
- Pinelle D, Nacenta M, Gutwin C, Stach T (2008) The effects of co-present embodiments on awareness and collaboration in tabletop groupware. In: Proceedings of GI 2008, GI '08, Canadian Information Processing Society, Toronto, Ont., Canada, pp 1–8
- Qvarfordt P, Beymer D, Zhai S (2005) Realtourist: a study of augmenting human–human and human–computer dialogue with eye-gaze overlay. In: Proceedings of INTERACT 2005, pp 767–780
- Qvarfordt P, Zhai S (2005) Conversing with the user based on eye-gaze patterns. In: Proceedings of CHI, ACM, pp 221–230
- Rogers Y, Lindley S (2004) Collaborating around vertical and horizontal large interactive displays: which way is best? *Interact Comput* 16(6):1133–1152
- Sellen AJ (1995) Remote conversations: the effects of mediating talk with technology. *Hum Comput Interact* 10(4):401–444
- Stein R, Brennan SE (2004) Another person's eye gaze as a cue in solving programming problems. In: Proceedings of ICMI, ACM, pp 9–15
- Steptoe W, Wolff R, Murgia A, Guimaraes E, Rae J, Sharkey P, Roberts D, Steed A (2008) Eye-tracking for avatar eye-gaze and interaction analysis in immersive collaborative virtual environments. In: Proceedings of CSCW 2008, ACM, pp 197–200
- Vertegaal R (1999) The gaze groupware system: Mediating joint attention in multiparty communication and collaboration. In: Proceedings of CHI, ACM, pp 294–301
- Wang H, Blevins E (2004) Concepts that support collocated collaborative work inspired by the specific context of industrial designers. In: Proceedings of CSCW, ACM, pp 546–549
- Whittaker S, O'Conaill B (1997) The role of vision in face-to-face and mediated communication, video-mediated

- communication ed. Lawrence Erlbaum Associates, ch. The role of vision in face-to-face and mediated communication, pp 23–49
30. Zhang Y, Bulling A, Gellersen H (2013) Sideways: A gaze interface for spontaneous interaction with situated displays. In: Proceedings of CHI, ACM Press, pp 851–860
 31. Zhang Y, Bulling A, Gellersen H (2014) Pupil-canthy-ratio: A calibration-free method for tracking horizontal gaze direction. In: Proceedings of AVI, ACM Press, pp 129–132
 32. Zhang Y, Chong MK, Müller J, Bulling A, Gellersen H (2015) Eye tracking for public displays in the wild. *Pers Ubiquitous Comput* 19(5–6):967–981
 33. Zhang Y, Müller HJ, Chong MK, Bulling A, Gellersen H (2014) Gazehorizon: Enabling passers-by to interact with public displays by gaze. In: Proceedings of UbiComp, pp 559–563
 34. Zhang Y, Stellmach S, Sellen A, Blake A (2015) The costs and benefits of combining gaze and hand gestures for remote interaction. In: Abascal J, Barbosa S, Fetter M, Gross T, Palanque P, Winckler M (eds) *Human-computer interaction*. Springer, Berlin, pp 570–577