Anticipation Without Acceleration: Benefits of Shared Gaze in Collocated Augmented Reality Collaboration

Julian Rasch
LMU Munich
Munich, Germany
julian.rasch@ifi.lmu.de

Vladislav Dmitrievic Rusakov
LMU Munich
Munich, Germany
vladislav.rusakov@campus.lmu.de

Jan Leusmann
LMU Munich
Munich, Germany
jan.leusmann@ifi.lmu.de

Florian Müller
TU Darmstadt
Darmstadt, Germany
florian.mueller@tu-darmstadt.de

Albrecht Schmidt
LMU Munich
Munich, Germany
albrecht.schmidt@ifi.lmu.de

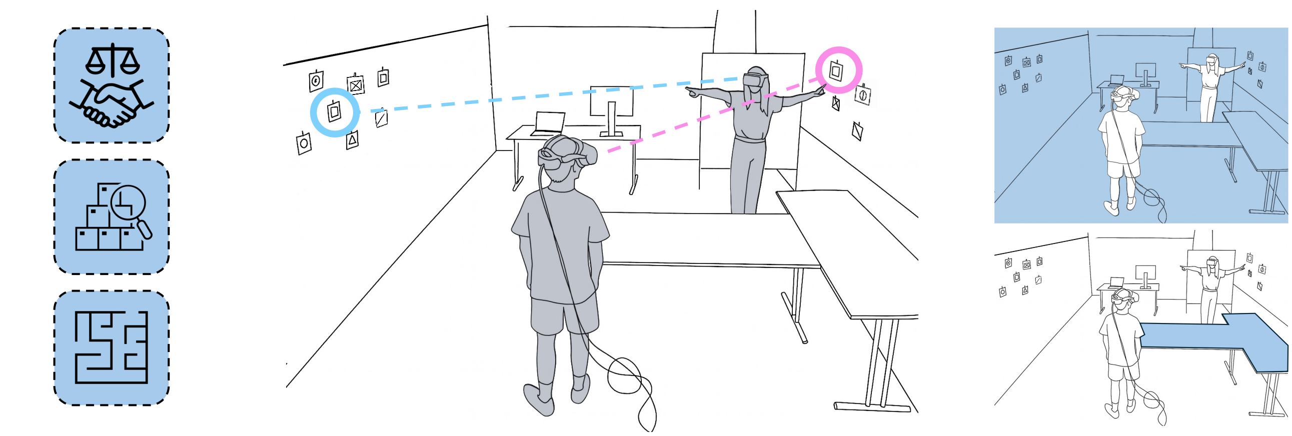


Figure 1: In this paper we study the effects of shared gaze (B) on user performance, social connectedness, and shared attention in collocated Augmented Reality (AR) collaboration across three distinct task types (A) in two interaction spaces (C).

Abstract

Knowing what collaborators attend to is essential. Previous studies demonstrated that shared gaze enhances coordination and social connectedness in remote settings. In collocated settings, gaze can be both naturally observable and technologically augmented. AR enables gaze cues to be rendered explicitly in the environment. To investigate if and how such cues are beneficial in collocated AR collaboration, we examined both qualitative and quantitative effects across three task types (puzzle, negotiation, search) and two spatial setups (plane, room), focusing on task completion time and the collaborative experience. In our user study with 24 dyads (n=48), we varied gaze visibility and measured task performance, user preference, social connectedness, and shared attention. Our results show that sharing gaze in collocated collaborative AR can increase shared attention, is perceived as helpful, and improves the user experience, similar to remote collaboration, but has a limited impact on the actual task completion time across the chosen tasks.

CCS Concepts

• **Human-centered computing** → **Interactive systems and tools**; **Interaction techniques**; **Mixed / augmented reality**; **Collaborative and social computing**.

Keywords

Multi-User, AR, XR, Co-Located, Joint Attention

CHI '26, Barcelona, Spain

ACM ISBN 979-8-4007-2278-3/26/04
https://doi.org/10.1145/3772318.3791758

ACM Reference Format:
Julian Rasch, Vladislav Dmitrievic Rusakov, Jan Leusmann, Florian Müller, and Albrecht Schmidt. 2026. Anticipation Without Acceleration: Benefits of Shared Gaze in Collocated Augmented Reality Collaboration. In *Proceedings of the 2026 CHI Conference on Human Factors in Computing Systems (CHI '26), April 13–17, 2026, Barcelona, Spain.* ACM, New York, NY, USA, 17 pages. https://doi.org/10.1145/3772318.3791758

1 Introduction

In face-to-face collaboration, gaze contributes to guiding and coordinating attention, disambiguating references, and establishing common ground [4, 5, 57]. However, natural gaze has limited spatial precision as the signal indicates a direction, not an addressable 3D target, degrades with distance and clutter, and requires visual attention to the sender. As a result, partners frequently need verbal or gestural support to establish joint attention and maintain smooth interaction. Augmented Reality (AR) can enhance collaboration by integrating digital cues into the physical space. This fundamentally changes how gaze can be utilized by making it explicit, persistent, and precisely located in shared space. This approach addresses several of the challenges outlined above: gaze no longer requires the recipient to look at the sender, and the cues can be spatially grounded with greater precision.

Prior work shows that shared gaze enhances remote collaboration by improving turn-taking, mutual awareness, and joint attention [4, 56, 58]. Studies further demonstrate that gaze cues support deictic referencing and the establishment of common ground, thereby strengthening coordination across distributed settings [12, 22, 41]. More recent work investigated gaze fixations in an asymmetric 2D–AR collaboration and found that collaborators could use gaze as a fast and precise pointer, often combined with other modalities [20]. Jing et al. [24] showed that in asymmetric remote AR–VR collaboration, bi-directional gaze cues amplified joint attention and improved co-presence compared to a no-gaze condition. They also extended findings from remote XR settings to collocated collaborative XR environments, suggesting that displaying shared gaze cues improves task performance compared to scenarios without gaze visibility in collocated contexts as well [23]. In collocated settings, making gaze explicitly visible sharpens references and sustains awareness beyond natural face-to-face cues [12, 22]. These works underline that shared gaze can improve coordination and foster social connectedness in both collocated 2D and XR collaboration. However, explicit gaze cues may also create challenges such as distraction [58], oversharing of attention, or inappropriate visibility, making their value highly dependent on spatial configuration and task context, highlighting the need to better understand how gaze sharing scales across different spaces and tasks.

Our work contributes to the body of literature on collocated gaze sharing by exploring the impact of gaze cues across varying tasks and collocated interaction spaces and addresses the following central research question: *"How does shared gaze affect social connectedness, cognitive load, task completion time, and shared attention across different task types and interaction spaces?"* The main goal of our study is to quantify the impact of seeing a collaborator's gaze in collocated settings across varying task types and interaction spaces. To isolate the specific collaborative value of gaze sharing from the general visual impact of an AR gaze cursor, we use a private gaze baseline that controls for visual distraction by displaying the participant's own gaze cursor.

We examine how different tasks and interaction spaces in collocated AR settings influence the impact of shared gaze cues, considering both objective outcomes and more subjective dimensions. We conducted a user study with 24 dyads (n = 48) performing three collaborative task types covering diverging collaboration styles (puzzle, negotiation, search), varying gaze visibility (only own gaze indicator, also other's gaze), and interaction space (table surface, room-scale). Here, we measured the effects of gaze visibility on social connectedness, distraction, performance, and mutual understanding. During the study, participants could either see their own gaze alone or additionally the other person's gaze. Participants worked on the collaborative tasks either on a table located between them (plane) or on the walls around the study space (room). This design allowed us to evaluate how shared gaze visualization affects coordination, social connectedness, shared attention, and user comfort across task settings.

Our results converge into five findings: (1) Collocated AR already affords natural joint attention through speech and gesture, so that shared gaze visualization does not add to social presence. This diverges from findings in remote and asymmetric settings where making gaze visible increased social presence / co-presence [3, 44]. (2) When displayed, gaze reliably steers attention, producing more joint fixations and stronger anticipation. This extends findings from previous remote studies, which show that shared gaze improves deictic reference, shared attention, and coordination in collaborative tasks [11, 23, 41]. (3) Shared gaze visualization does not alter perceived workload in the tasks at hand. While some previous work reports reduced effort in spatially distributed tasks [3, 24], others report increased distraction in simple or local tasks [58] along with higher physical and comparable mental demand [58]. Overall, this highlights that the effects of shared gaze on effort remain mixed and are likely influenced by differences in tasks, measures, and system design. (4) Shared gaze does not improve efficiency in our study context. This contrasts prior findings that shared gaze can shorten task completion time and improve performance in collaborative scenarios [3, 17, 23, 41, 58] but also connects with some other findings [11]. (5) Users prefer having shared gaze, especially for spatial problem-solving tasks, even without objective gains in task completion time. This preference is consistent with earlier findings that users tend to like or choose systems with shared gaze cues [17, 23, 44] and, together with additional attention steering (Finding (2)) and the lack of differences in workload (Finding (3)), suggests that people may still prefer shared gaze even when it does not lead to measurable performance gains.

Besides these data-driven findings, based on participants' responses after the study, we also identified concerns that participants have about potentially using a gaze-sharing system in scenarios from their day-to-day lives. These findings provide guidance on when shared gaze is advantageous in AR collaboration and inform the design of future systems.

2 Related Work

Gaze is a crucial communication cue in collaboration, providing a fast and precise way to coordinate and establish common references. It can function as a pointer [20], confirm or clarify the object of interest [13, 53], and simplify complex referring expressions such as *"this piece over here"* [34]. Researchers have also identified a gaze leading effect, a social orienting response in which a person's attention is drawn to the face of a collaborator who has followed their gaze to an object [14]. This mutual gaze awareness further supports joint attention, enabling collaborators to reach shared understanding and effective decision-making [58]. Psycholinguistic research underscores this collaborative function: Clark and Wilkes-Gibbs [9] showed that reference is established not by a one-sided description but through iterative alignment between partners. Their tangram studies revealed how shared attention allows collaborators to converge on shorter, more efficient expressions, and highlight gaze as one resource for grounding communication. Besides social functions, gaze is also well explored as an input modality for 2D screens as well as XR settings [45, 51]. Beyond gaze as input, previous works also explored the potential of (explicitly) sharing gaze through technology with a collaborator. In the following, we distinguish prior work on shared gaze in *remote collaboration* and in *collocated collaboration*.

2.1 Shared Gaze in Remote Collaboration

In remote collaboration, sharing gaze through eye-tracking and visualizations provides benefits for coordination. Seeing a collaborator's gaze allows partners to monitor each other's focus of attention [33] and improves mutual awareness when body position and head orientation are not visible. Ishii and Kobayashi [22] introduced the importance of "gaze awareness" for smoothly coordinating talk and drawing, highlighting how visibility of the partner's eyes supports collaborative problem solving. Shared gaze further enhances co-presence even when physically apart [17]. Prior work shows that mutual gaze facilitates joint attention [11, 24], supports deictic referencing [1], and contributes to establishing common ground in distributed decision-making [41, 56]. Shared gaze can also enable remote collaborators to adopt coordination strategies known from collocated interaction (e.g., implicit deixis) [12]. To make gaze visible, researchers have introduced a variety of visualization styles, including cursors/dots, spotlights, trails/paths, and heatmaps [4, 13, 42]. These representations can help separate the task space (e.g., divide areas for visual search) [4] or reveal a partner's intentions [42]. At the same time, overlays can impose costs if poorly chosen or error-prone: continuous trails may clutter workspaces, whereas shared-area cues that only highlight overlapping attention can support coordination with less distraction [13]. Moreover, in AR, degraded accuracy/latency in shared gaze cues can reduce performance and affect user experience, underscoring the need to respect error budgets [15]. Evidence on flat displays is extensive [46]: shared gaze improves coordination in visual search [4], writing [32], and programming [11, 53]. In education, it boosts learning gains and helps instructors track understanding [55], while in robotic surgery, it enhances coordination between operators [31]. Shared gaze for remote collaboration has also been explored for symmetric and asymmetric immersive modes. AR-2D and video-capture head-mounted display (HMD)-2D systems project or overlay fixations into the physical workspace, providing a fast and precise pointer that complements gestures and speech [1, 17, 20]. AR–VR systems share gaze alongside head/hand cues, and studies report that eye gaze offers high precision for object referencing while head gaze and field-of-view cues are coarser but often more interpretable [44]. Bi-directional gaze in MR strengthened co-presence and reduced verbal effort (e.g., HoloLens/VR) [3, 24, 25], and combining gaze with natural gestures further lowered workload in remote assistance [48]. Recent work also compared communication modalities in VR (head gaze, gesture, and voice, together with their combinations) and found that, despite user preference in favor of all three modalities combined, this combination did not perform significantly better than any other combination or head gaze alone [16]. However, voice or gesture alone performed significantly worse.

Beyond synchronous dyads, asynchronous gaze cues have shown mixed outcomes: while gaze did not consistently improve objective performance, some participants preferred its availability [49]. Reviews of multiparty collaboration emphasize that gaze metrics (e.g., joint attention) are diagnostic of group dynamics and have been associated with improved cooperation and learning [46].

In sum, there is broad evidence that remote shared gaze reliably improves coordination, joint attention, and co-presence, with task- and design-dependent performance gains. However, research also found that overlays impose costs when inaccurate or visually heavy.

2.2 Shared Gaze in Collocated Collaboration

Making gaze explicit through visualization turns implicit attention into spatially precise cues. On large 2D displays, showing others' gaze can accelerate collaborative search and ease communication, but introduces trade-offs in distraction and privacy [58]. Multi-view tabletops extend this by letting collaborators shift between private and shared views while maintaining awareness of others' focus [37]. Projection-based systems further externalize group attention without head-worn devices by visualizing multiple users' gaze on a shared surface, supporting small-group coordination [29]. Hybrid configurations combine a public display with optical see-through HMD overlays, aiming to preserve natural eye contact and gestures while adding private views that can be selectively shared [54]. Display choices matter for retaining interpersonal cues: optical see-through AR preserves face-to-face gaze/gesture more effectively than immersive VR, which can prompt compensatory speech [28].

Within AR HMDs, comparisons of gaze visualizations reveal consistent trade-offs. Head-/eye-based rays support efficient, accurate reference identification, whereas trailing cues (e.g., moving tracks) are rated higher for social presence and preference [7]. For wide-area, model-free AR, bracketing with *Double Ray* substantially reduces visual ambiguity, while *Parallel Bars* add cognitive effort without accuracy gains [35]. Bi-directional gaze in collocated AR reinforces mutual awareness: *Laser Eye* was preferred and associated with lower effort, while *Trail Path* helped with history yet risked distraction in tight workspaces [23].

Beyond coordination, joint attention dynamics relate to learning outcomes: groups that achieved stronger shared attention showed higher learning gains, whereas low-achieving groups often failed to converge visually [27]. In educational use, immersive gaze sharing helped instructors assess learners' focus and comprehension, but some learners reported discomfort from feeling watched, indicating acceptance depends on activity and privacy controls [26]. Synthesis of literature indicates that while shared gaze often benefits collaboration and co-presence in XR, effectively realizing these benefits for collocated immersive settings remains an open challenge [19].

Taken together, prior work highlights the promise of shared gaze for enhancing coordination, awareness, and social connection in collocated collaboration, but evidence of consistent performance benefits is mixed.

2.3 Summary and Research Gap

Prior work has established the value of gaze for collaboration in both 2D screen-based and XR settings for remote and collocated collaboration. Findings suggest performance gains when shared gaze is made explicitly visible in remote as well as collocated contexts, yet clear evidence of actual performance improvements in collocated settings remains limited. In remote scenarios, gaze cues help to communicate information that is not directly available. In collocated situations on a tabletop/in 2D, that same cue can be redundant or even distracting because gaze and gestures are visible naturally. Across room-scale, with distance and occlusion increasing, the same cue can quickly establish joint attention without extra pointing. The same is true for different tasks: While spatially distributed search tasks might benefit from a precisely anchored gaze indicator, the same indicator might be distracting in other tasks.

Building on prior work reviewed across remote and collocated collaboration, we therefore investigate the impact of shared gaze in symmetric collocated XR collaboration on (1) shared attention, (2) verbal communication, (3) social connectedness, (4) perceived helpfulness, (5) task performance, and (6) user preference. Based on these considerations, we formulate the following hypotheses.

H1: Sharing gaze with a collocated collaborator reduces task completion time compared to collaboration without gaze sharing [3, 17, 23, 41, 58].
H2: Sharing gaze with a collocated collaborator increases shared attention [11, 23, 41].
H3: Sharing gaze with a collocated collaborator increases social connectedness [3, 44].
H4: Sharing gaze with a collocated collaborator is preferred by users [17, 23, 44].
H5: Sharing gaze with a collocated collaborator reduces users' effort [3, 24] in spatially distributed tasks, but increases distraction in simple or local tasks [23, 26, 58].
H6: The benefits of sharing gaze with a collocated collaborator vary across task types and interaction space [26, 58].

3 Methodology

To study the effects of shared gaze visibility in collocated AR collaboration across different task settings, we selected tasks from three categories commonly used in gaze-supported collaboration studies: puzzle [24, 43], negotiation [58], and visual search [4, 23]. We chose each task for its applicability in both close and far proximity, for showing potential benefits of shared gaze, and for requiring symmetric involvement of two users. All tasks are grounded in real-world interactions, with AR used solely to display gaze cues. We used a within-subject design, varying three independent variables (IV). The first two IVs define the task setting: Task Type (Puzzle / Negotiation / Search) and Interaction Space (Plane: a close 2D surface / Room: the whole 3D study space). In each setting, we further vary Gaze Visibility between Private Gaze and Shared Gaze. This results in a $2 \times 2 \times 3$ design with 12 conditions.

To minimize order effects, conditions were arranged in a three-level nested counterbalancing scheme, corresponding directly to the three IVs: 1) Gaze Visibility: Participants experienced conditions in blocks defined by Gaze Visibility. The two visibility levels (Private Gaze vs. Shared Gaze) were organized using a Latin square across participant groups. 2) Interaction Space: Within each visibility block, the two Interaction Space levels (Plane vs. Room) were counterbalanced so that both occurred equally often in each order position across participants. 3) Task Type: Within each interaction space block, participants completed the three Task Types (Puzzle / Negotiation / Search). The task order was fully counterbalanced using all six possible permutations, ensuring that each task appeared equally often in each ordinal position. This hierarchical structure resulted in 24 unique sequences, which were distributed evenly across groups of participants. Thus, order and fatigue effects were controlled systematically at the block level (Gaze Visibility), the sub-block level (Interaction Space), and the within-condition level (Task Type). To further avoid learning effects of the task itself, since it was repeated twice, we prepared multiple versions of each task. For Negotiation, this was achieved by rearranging the task feature cards on the interaction space as well as giving participants different pairs of conflicting task goals. For the Puzzle we prepared multiple path puzzles, and for Search we changed the pairs and rearranged the pieces. These approaches help minimize learning and order effects of the tasks themselves.

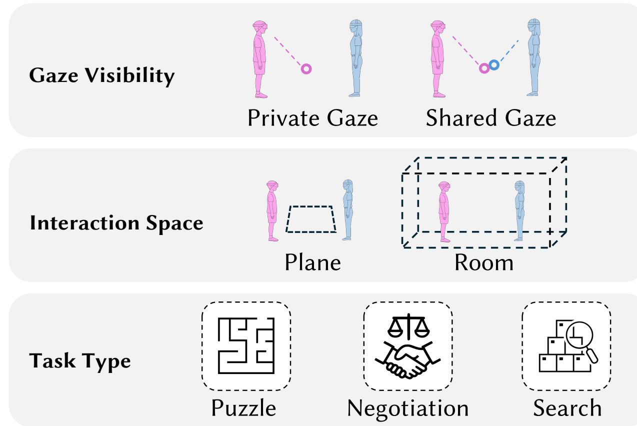


Figure 2: The three Independent Variables of our study: Gaze Visibility, Interaction Space, and Task Type with their respective levels.

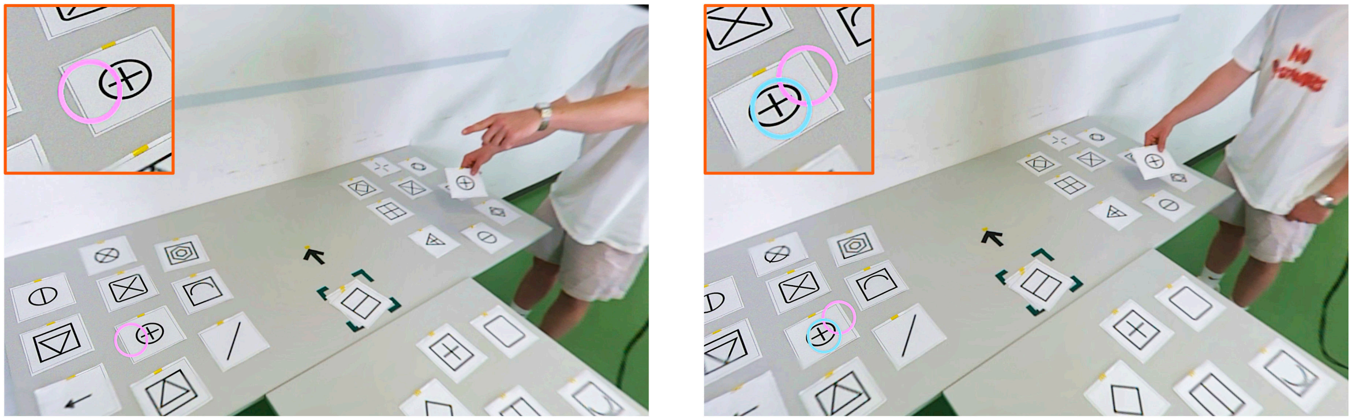
3.1 Independent Variables

We vary the following three independent variables as illustrated in Figure 2. As we are interested in symmetric collaboration, in each condition, both participants experience the same level of each IV.

Gaze Visibility We vary the Gaze Visibility on two levels to be either displayed as Private Gaze or Shared Gaze. We display the gaze position in the form of an O-shaped cursor in pink (own) and blue (other), similar to closely connected previous works [23, 24]. Figure 3 shows this implementation as seen by the participants.

Private Gaze: Participants only see their own gaze.

Shared Gaze: Participants see own and other person's gaze. Although related works on single-user gaze visualization usually do not display a user's own gaze cursor, multi-user systems that provide shared gaze typically show both users' gaze pointers [25]. In line with this, Shared Gaze displays both the user's own



(a) Private Gaze **(b) Shared Gaze**

Figure 3: The two levels of Gaze Visibility during the Plane – Search Task: (a) Private Gaze showing only the participant's own gaze cursor as a pink circle, and (b) Shared Gaze additionally showing the collaborator's gaze cursor in blue, as seen from the participant's perspective. The top-left inset in each image shows a close-up for better visibility.

and the collaborator's gaze pointers. Private Gaze serves as a baseline that only shows the own gaze. This baseline allows us to isolate the specific contribution of the collaborator's gaze while keeping the generic visual and interaction effects of having a gaze cursor on the screen (e.g., an additional moving visual element and potential distraction) constant. A baseline with no gaze cursor would change two factors at once, namely the presence of a gaze cursor at all and the presence of the collaborator's gaze, which would make it difficult to attribute any performance or experience differences uniquely to shared gaze. Our focus is therefore on the incremental effect of shared gaze cues in a collaborative setting, rather than on reflecting an accurate single-user setting without gaze cursors.

Interaction Space Participants collaborate on the tasks in two different Interaction Spaces. We chose those based on interaction spaces common in collocated collaboration in office-like environments. Figure 4 shows the Task Types for both levels of the Interaction Spaces.

Plane: A 2D table surface in close proximity, positioned in between the two participants.

Room: The whole 3D space of the study room. Here, in contrast to Plane, the task space and task pieces are not only in close proximity but distributed around the space.

Task Type We chose three Task Types representative of collaborative collocated tasks adapted from related works to identify tasks that benefit most from sharing gaze. Figure 4 shows the Task Types for both levels of the Interaction Spaces.

Puzzle: The path puzzle is a structured collaborative task in which two participants jointly construct a continuous path on a grid using a fixed set of straight and curved pieces. Each turn, one participant instructs the other which piece to place and where, after which roles switch. Errors may be corrected by replacing pieces following the same turn-taking procedure. The task concludes once all pieces are placed and a complete path from the predefined start to end point is achieved. Similar to [24].

Negotiation: The negotiation task places two participants in a decision-making scenario where they must jointly select a hotel from a set of options. Each hotel is represented on a "map" with the rated attributes *cleanliness, location, service, facilities, price, comfort, and wifi*. We designed the hotels to have conflicting qualities (cleanliness vs. wifi, comfort vs. service, facilities vs. price, price vs. location). While both participants share the overarching goal of staying close to parks and avoiding nearby construction sites, their individual preferences conflict (e.g., one prioritizes low cost, the other prioritizes cleanliness). This creates a structured but open-ended negotiation problem: participants must weigh trade-offs, justify their positions, and reach an agreement despite misaligned priorities. Similar to [58].

Search: The visual search task requires two participants to collaboratively identify and match pairs of geometric shapes printed on sheets of paper. The goal is to locate five correct pairs within 40 pieces and physically place them onto a designated area on a desk. This setup creates a shared visual search problem in which participants must coordinate attention, compare alternatives, and reach agreement on matching items. The search task is designed to create joint attentional processes, coordination strategies, and a division of labor in a controlled but interaction-rich setting. Similar to [4, 23].

3.2 Dependent Variables

During each condition, we collect participants' data in the form of log data and questionnaire responses after the condition. We employ three established questionnaires as well as custom questions after each condition. After completing all conditions, we further collect participants' responses via a post-study survey.

IOS Score: We use the Inclusion of Other in the Self (IOS) questionnaire [2] as a measure for social closeness of the two participants. The IOS is a single-item questionnaire with a 7-point answer scale. We use the IOS score to assess H3.

GEQ Score: We use the Behavioural Involvement Component in the Social Presence Module of the Game Experience Questionnaire

(a) Plane (b) Room

Figure 4: Task setup for the three Task Types (Puzzle, Negotiation, and Search) across the two Interaction Spaces.

(GEQ) [21] as a measure for social presence and shared attention. The component consists of 6 items on a 5-point Likert scale (not at all, slightly, moderately, fairly, extremely). We use the GEQ Score to assess H2 and H3.

RAW-TLX Score: We use the NASA Task Load Index (TLX) as a measure for task demand and subjective performance rating, consisting of 6 subscales. The subscales are Physical Demand, Mental Demand, Temporal Demand, Performance, Effort, and Frustration. We calculate the Raw Nasa-TLX (RTLX) as suggested by Hart [18] as the mean over all subscales without weights. Participants answered each item on a 21-point answer scale, representing a 0-100 scale in steps of 5. We use the RTLX to assess H5 and H6.

Custom Questions: We further administered a custom questionnaire with items regarding one's own and the other's communication cues to assess how helpful participants perceive them. Participants rated both their own and the other's communication cues on anchored 5-point Likert scales (1 = strongly disagree to 5 = strongly agree). We adapted questions for the first 5 items [47] to assess if cues were understandable, effective, intrusive, disturbing, and noticeable. We additionally added questions regarding their role in reducing misunderstandings and enabling anticipation of next actions. Each dimension was measured symmetrically for self- and other-directed communication cues. Participants also rated whether they perceived their own as well as the other's gaze cues as distracting and helpful. For conditions in which no other's gaze cues were visible, we instructed participants to indicate any answer, as we analyzed this data only for the Shared Gaze conditions. We also included an open comment field for participants' text feedback after each condition. We use these measures to assess H2, H4, H5, and H6.

Post-Study Questionnaire: After completing all conditions, participants further complete a post-study questionnaire, holding custom questions regarding preference of a shared gaze system for solving puzzle, negotiation, or search tasks in the future on anchored 5-point Likert scales (1 = strongly disagree to 5 = strongly agree). We use this data to assess H4 and H6.

Qualitative Data: To collect qualitative data going beyond the scope of the RQ of this study, we further let participants fill out open text fields regarding situations they would like to use a shared gaze system in, situations where they might be hesitant, and inquired about explanations of these situations (see Appendix A). We chose this approach over an interview due to the already long study time and the dyadic setup, as it was faster but still allowed us to gain deeper insights into the participants' thoughts.

Besides the subjective measure, we further logged and recorded participants' interactions during each condition. Here, we logged gaze data and task completion times.

Gaze Data: During each condition, we logged participants' eye gaze data using the built-in eye tracker of the Varjo XR-4 Focal Edition[1]. We use this data to assess H2.

Task Completion Time: We logged participants' Task Completion Time for each condition. We calculate the Task Completion Time as the difference between the UNIX time at task completion and the UNIX time stamp at task start. We use this data as the primary performance measure and to assess H1.

Audio Recordings: We further recorded audio for each participant using the directional microphones of each HMD. We combine the audio tracks and transcribe the audio recording to count words to identify changes in verbal communication.

3.3 Apparatus

We developed the application in Unity[2](Version 2022.3.57f1). We used two tower PCs suitable for the Varjo XR-4 Focal Edition HMDs. PC1: Intel Core i9-13900KF (3.00 GHz), 64 GB RAM, Windows 11 Pro, NVIDIA GeForce RTX 4090 (24 GB). PC2: Intel Core i9-12900K (3.20 GHz), 64 GB RAM, Windows 11 Pro, NVIDIA GeForce RTX 3090 Ti (24 GB). We used two Varjo XR-4 Focal Edition HMDs with a 120° x 105° field of view and a resolution of 2840 x 2816 (peripheral area) and 1680 x 1680 (focus area) pixels (foveated 55 PPD) with 90 fps. We used the built-in eye tracker to collect participants' gaze data and logged eye and gaze position 90 times per second locally on each computer. We further tracked participants' positions with 4 SteamVR Base Stations 2.0[3]. To realize the collaborative setup, we transmitted their gaze point and position data via network using Photon PUN 2[4]. We provided an overall interaction space of approximately $4m \times 4.5m$. Depending on the condition, we asked participants to collaborate on one of three Task Types. Most relevant were the two opposing walls, a movable whiteboard, two tables arranged in a T shape, positioned as indicated in Figure 5.

3.4 Procedure

After completing the informed consent and answering potential questions, participants began the demographics survey. After this, we explained the task for the first condition and let them put on the HMDs. Before each condition, we calibrated the Varjo XR-4

[1]https://varjo.com/products/xr-4/
[2]https://unity.com/
[3]https://www.vive.com/
[4]https://www.photonengine.com/pun

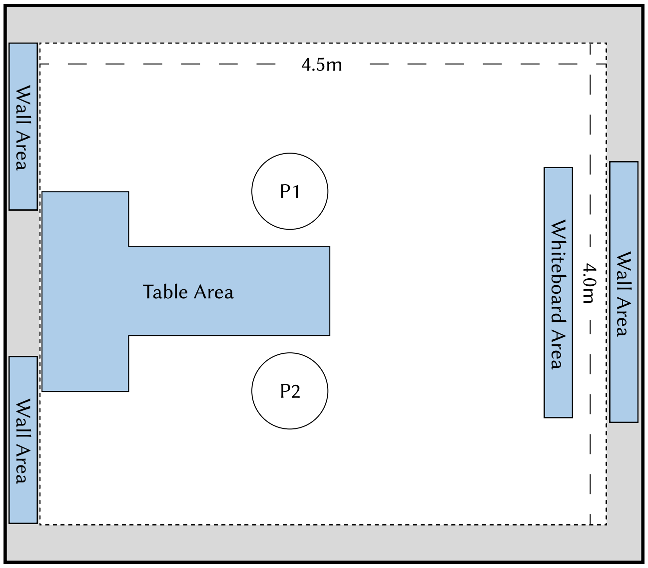


Figure 5: Top view of the study area. PLANE tasks occurred on the tables, while ROOM tasks additionally occurred on the walls (NEGOTIATION, SEARCH) and whiteboard (PUZZLE).

built-in eye tracker using its 5-point calibration process. After this, participants saw their own gaze in the form of an O-shaped cursor. We used a cross on one wall of the room to let participants verify that the tracking is accurate. If not, we repeated the calibration process until participants confirmed an accurate representation of their gaze. After this, the experimenter again explained the task. Once participants confirmed they were ready, we started the condition together with the condition timer. Upon completion of the task, participants informed the experimenter that they were finished, and we stopped the timer. Participants removed their HMDs and sat down at a laptop each to complete a survey after each condition. After this, participants put their HMDs back on, followed by the same calibration process as before, before continuing with the next condition. After completing all 12 conditions, participants filled out an additional survey before concluding the study. Participants received compensation for their participation of 6€ per 30 minutes. On average, the study lasted approximately 120 minutes, including breaks.

3.5 Participants

We recruited 48 participants through our university's email service and printouts. Before starting the experiment, we collected participants' demographic data and familiarity data of the pairs. 23 participants identified as male and 25 as female. Participants' mean age was 27.3 years, ranging from 21 to 60 years. Looking at mean values per pair of participants' answers regarding the statement *"I know the other person very well"* on a 7-point anchored Likert scale (strongly disagree to strongly agree), 9 pairs of participants rated their familiarity high (5-7), and 15 low (1-3), with no neutral ratings. When asked *"How often did you experience Augmented Reality on a Head Mount Display before?"* 9 participants stated never, 27 stated 1-5 times, 5 stated 6-10 times, and 7 stated 10+ times.

4 Results

We report the results of our analysis of objective log data, subjective ratings from questionnaires, and qualitative data based on participants' text responses.

4.1 Subjective Measures

Following the ongoing debate around the appropriateness of ART and similar approaches for analyzing ordinal data sets [36, 39], we opted for ordinal regression models in the analysis following established guidelines from psychology [6]. For this, we used cumulative link mixed models (CLMMs) fitted with the `ordinal` (R) package. We included GAZE VISIBILITY, INTERACTION SPACE, TASK TYPE, and all interactions as fixed effects. Further, we added random intercepts for participants and groups and per-participant random slopes where convergence allowed. Following common practice [50], we analyzed averaged values for multi-item scales such as NASA-TLX and GEQ with LMEs treating the aggregate as approximately continuous, similar to the task-completion time.

While we provide all results of our statistical analysis for the subjective ratings in Table 1 and Table 2, we will focus only on effects of GAZE VISIBILITY as our central IV in this subsection. INTERACTION SPACE and TASK TYPE are treated as covariates, as they are only relevant to our research question and hypotheses in combination with GAZE VISIBILITY.

4.1.1 IOS. We analyzed the IOS score as a measure of social closeness of the two participants. We found no significant main nor interaction effects of GAZE VISIBILITY on the IOS score.

4.1.2 GEQ. We calculated the GEQ score as the average of the items of the Behavioral Involvement Component of the social presence module of the GEQ, according to the scoring guidelines [21]. Our analysis showed no significant main effect of GAZE VISIBILITY nor significant interaction effects involving GAZE VISIBILITY.

4.1.3 RTLX. For the RTLX score, we calculated the mean of the six subscales without weights as suggested by Hart [18]. We found no significant main effect for GAZE VISIBILITY nor significant interaction effects involving GAZE VISIBILITY. Significant effects for TASK TYPE and INTERACTION SPACE as well as significant interaction effects are shown in Table 1.

4.1.4 Custom Questionnaire. In this subsection, we report the results of our custom questionnaire (Table 2). The exact wording of all questions is provided in Table 3 in the Appendix A.

Other's Communication Cues. We found no significant main effect for GAZE VISIBILITY nor significant interaction effects on participants' ratings regarding *CQ1-5: "The other's communication cues were very..." a) "understandable", b) "effective", c) "intrusive", d) "disturbing", e) "noticeable"* (see Table 2). Also for *CQ6: "The other's communication cues reduced misunderstandings a lot."* we found no significant main effect for GAZE VISIBILITY nor significant interaction effects on participants' ratings. However for *CQ7: "I could anticipate the others' next action very well."*, we found a significant main effect for GAZE VISIBILITY on participants' ratings, with significantly higher ratings for SHARED GAZE compared to PRIVATE GAZE. We further found a significant interaction effect for GAZE VISIBILITY × TASK TYPE.

Table 1: ANOVA results for GEQ and RTLX. Task Type (T), Gaze Visibility (G), and Interaction Space (S).

	T		G		S		T×G		T×S		G×S		T×G×S	
Variable	*F*	p	*F*	p	*F*	p	*F*	p	*F*	p	*F*	p	*F*	p
GEQ	34.58	<.001	1.66	.204	2.49	.122	1.46	.233	1.21	.299	1.76	.186	0.78	.459
RTLX	50.61	<.001	1.40	.243	14.86	<.001	0.50	.608	8.23	<.001	0.68	.412	1.00	.370

Table 2: CLMM results with Wald χ^2 tests for the IOS score and for all items CQ1-CQ18 of the custom questionnaire (see Table 3). Task Type (T), Gaze Visibility (G), and Interaction Space (S).

	T		G		S		T×G		T×S		G×S		T×G×S	
Variable	$\chi^2(2)$	p	$\chi^2(1)$	p	$\chi^2(1)$	p	$\chi^2(2)$	p	$\chi^2(2)$	p	$\chi^2(1)$	p	$\chi^2(2)$	p
IOS	6.092	.048	3.457	.063	0.188	.665	3.662	.160	4.766	.092	0.857	.355	1.284	.526
Other's Cues														
CQ1: Understandable	2.136	.344	.827	.363	.637	.425	2.648	.266	.808	.668	.313	.576	2.208	.331
CQ2: Effective	.382	.826	.251	.616	.322	.571	.218	.897	1.178	.555	.083	.773	2.698	.260
CQ3: Intrusive	1.206	.547	1.901	.168	.015	.901	2.120	.346	.500	.779	2.442	.118	.302	.860
CQ4: Disturbing	1.262	.532	.113	.737	.120	.729	2.276	.320	1.158	.560	.806	.369	1.348	.510
CQ5: Noticeable	3.956	.138	1.249	.264	.002	.967	1.722	.423	.162	.922	.009	.925	4.382	.112
CQ6: Reduced Misunderstandings	3.366	.186	1.808	.179	1.665	.197	.782	.676	.510	.775	1.442	.230	.498	.780
CQ7: Anticipate Next Action	7.484	.024	4.569	.033	.431	.512	6.164	.046	.680	.712	.127	.722	.364	.833
Own Cues														
CQ8: Understandable	3.488	.175	1.469	.226	2.025	.155	1.256	.534	.246	.885	.398	.528	6.270	.044
CQ9: Effective	.304	.859	1.021	.312	.064	.800	1.514	.469	1.078	.583	.072	.788	.976	.614
CQ10: Intrusive	.864	.649	.195	.659	2.053	.152	2.814	.245	3.902	.142	.008	.927	3.300	.192
CQ11: Disturbing	3.794	.150	.354	.552	.476	.490	2.916	.233	1.904	.386	.224	.636	.470	.791
CQ12: Noticeable	2.026	.363	3.484	.062	.088	.767	.746	.688	.902	.637	4.029	.045	1.402	.496
CQ13: Reduced Misunderstandings	.494	.781	3.874	.049	.126	.723	1.416	.493	.062	.970	.499	.480	.922	.631
CQ14: Anticipate Next Action	8.578	.014	4.256	.039	.949	.330	11.044	.004	.744	.689	.954	.329	1.924	.382
Own Gaze was														
CQ15: Distracting	3.108	.212	8.090	.005	.578	.447	.368	.832	2.824	.244	2.764	.096	.296	.862
CQ16: Helpful	18.978	<.001	37.780	<.001	3.526	.060	2.308	.316	3.234	.199	.016	.901	1.480	.477
Other's Gaze was														
CQ17: Distracting	1.008	.604			1.847	.174			4.696	.096				
CQ18: Helpful	9.330	.009			4.164	.041			3.088	.214				

Own Communication Cues. We found no significant main effect for Gaze Visibility on participants' ratings regarding the statements *CQ8-12 "I felt my own communication cues were very..."* a) *"understandable"*, b) *"effective"*, c) *"intrusive"*, d) *"disturbing"*, e) *"noticeable"* (see Table 2). While we did not find significant interaction effects for own *effective, intrusive, disturbing* ratings, we found a significant interaction effect for Gaze Visibility × Task Type × Interaction Space on own:understandable. We further found a significant interaction effect for Gaze Visibility × Interaction Space on own:noticeable. For *CQ13: "I felt my own communication cues reduced misunderstandings a lot."* we found a significant main effect for Gaze Visibility on participants' ratings, with higher ratings for Shared Gaze compared to Private Gaze. We did not find any significant interaction effects involving Gaze Visibility. For *CQ14: "I felt the other could anticipate my next action very well."* we found a significant main effect for Gaze Visibility on participants' ratings, with higher ratings for Shared Gaze compared to Private Gaze. We further found a significant interaction effect for Gaze Visibility × Task Type.

Own Gaze. For the following items, we report effects for all IVs, as the question itself includes gaze. For *CQ15: "I found it very distracting seeing my own gaze visualized."* we found a significant main effect for Gaze Visibility on participants' ratings, with higher ratings for Private Gaze compared to Shared Gaze. We found no significant interaction effects. For *CQ16: "I found it very helpful seeing my own gaze visualized."* we found a significant main effect for Gaze Visibility on participants' ratings, with higher ratings for Shared Gaze compared to Private Gaze. We further found a significant main effect for Task Type on participants' ratings, with significantly higher ratings for Search and Puzzle compared to Negotiation. We found no significant interaction effects.

Other's Gaze. We only analyzed the following two items for Shared Gaze conditions, hence do not treat Gaze Visibility as an IV for the analysis. For *CQ17: "I found it very distracting seeing the others' gaze visualized."* we found no significant main effect for Task Type and Interaction Space on participants' ratings, nor significant interaction effects. However, for *CQ18: "I found it very helpful seeing the others' gaze visualized."* we found a significant main effect for Task Type on participants' ratings. Post-hoc tests confirmed significantly higher ratings for Search compared to Negotiation. We further found a significant main effect for Interaction Space on participants' ratings, with significantly higher ratings for Room compared to Plane. We found no significant interaction effects.

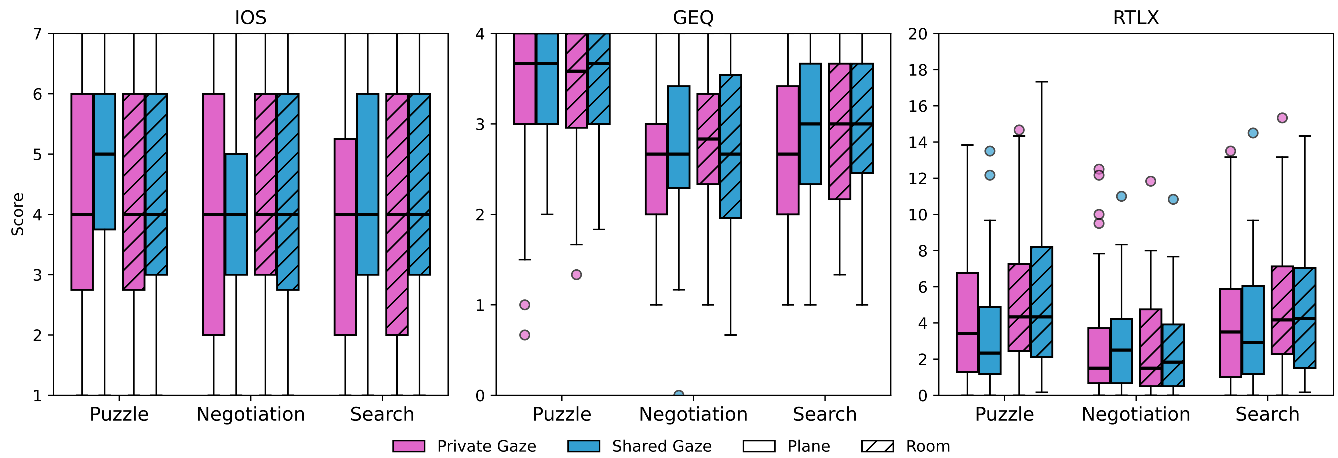


Figure 6: Results of the IOS, GEQ, and RTLX scores, visualized as boxplots with indicated median.

4.2 Task Completion Time

To assess task performance, we measured the duration of each task across conditions. The experimenter manually logged the start and end times (UNIX timestamps) for each task by pressing the *start* (S) and *end* (E) keys in the Unity application, respectively, at the beginning and completion of each task. Figure 9 depicts the distribution of the task completion time across all three levels of Task Type, two levels of Interaction Space, and two levels of Gaze Visibility. We analyzed the task completion time with linear mixed effects models (LMEs) on `log`-transformed durations to address a right-skew in the data. A maximal random-effects structure led to singular fits. The best converged model included a random intercept for `group` and a random slope for Gaze Visibility together with fixed effects for the three independent variables and their interactions. This model slightly outperformed the intercept-only variant by AIC, and both yielded consistent inferences. Post-hoc comparisons used `emmeans` with Bonferroni correction and results are reported on the response scale.

Wald F-tests with Kenward–Roger degrees of freedom showed a significant main effect of Task Type ($F_{2,230} = 139.6$, $p < .001$ with task completion time increasing from Negotiation ($M = 107.00$, $SD = 71.60$) over Search ($M = 194.00$, $SD = 86.60$) to Puzzle ($M = 333.00$, $SD = 218.00$), all $p < .001$. Further, the analysis indicated a significant main effect of Interaction Space ($F_{1,230} = 12.27$, $p < .001$) where Plane ($M = 176.00$, $SD = 125.00$) was faster than Room ($M = 247.00$, $SD = 199.00$). Finally, we found a Task Type × Interaction Space interaction ($F_{2,230} = 15.96$, $p < .001$) that moderated these main effects. Plane was faster than Room for Puzzle (Plane $M = 258.00$, $SD = 168.00$ vs. Room $M = 408.00$, $SD = 238.00$, $p < .001$) and Search (Plane $M = 154.00$, $SD = 48.00$ vs. Room $M = 235.00$, $SD = 97.00$, $p < .001$). For Negotiation, however, Room was faster than Plane (Room $M = 96.00$, $SD = 64.00$ vs. Plane $M = 117.00$, $SD = 78.00$, $p < .05$). Interestingly, Gaze Visibility and all interactions involving it were not significant (all $p > .38$).

In summary, time strongly depends on Task Type and Interaction Space, with a Task Type × Interaction Space pattern (Room advantage for Negotiation, Plane advantages for Puzzle and Search), while Gaze Visibility does not measurably affect task completion time.

4.3 Analysis of Gaze Data

To better understand how shared gaze influenced collaboration, we analyzed the recorded eye-tracking data to quantify joint attention. We sampled eye gaze data at 90Hz, using Unity's FixedUpdate loop, synchronized with the Varjo XR-4s display refresh rate. We logged both the target position where the gaze would be visualized and the eye positions, from which we could calculate the head position as the midpoint between the left and right eye positions. We then identified fixations using the Identification of Dispersion-Threshold (I-DT), implemented through the I-DTVR package [38]. Based on the suggestions of Llanes-Jurado et al. [38], we set the temporal threshold to 0.25s, the dispersion threshold to 1.6, and a minimum frequency threshold to 30Hz. This resulted in 122536 fixations. We then filtered out fixations that occurred outside the task intervals, as defined in Section 4.2, and excluded fixations outside each task's designated task areas to ensure we included only task-relevant gaze behavior. Lastly, to improve fixation accuracy, we merged temporal and spatially close micro-fixations using a 150ms temporal threshold and 5cm spatial threshold. With this, we reached our final analysis fixation count of 120983. To understand the offset in the gaze visualization between the two participants, we calculated the distance between each participant's target position and a fixed calibration point in the room, which both participants had to focus on before each task to verify their eye-tracking calibration. We then calculated the mean position of all gaze points close to that calibration point for both participants in a dyad before each task. We then averaged over all these positions to get an approximate offset between the visualizations ($M = .0599m, SD = .0415m$).

To compute joint attention, we aligned the fixation data of both participants in a dyad temporally and classified *joint attention* if the centroids of two fixations are below 25cm away from each other, while their time intervals overlap. We then calculated the ratio of joint attention fixations to the total number of fixations

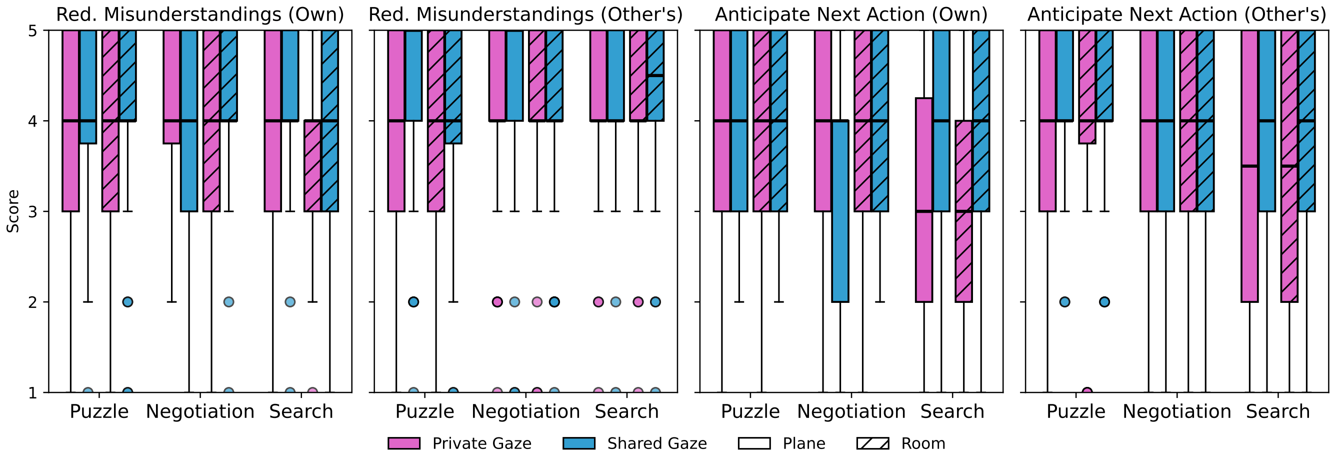


Figure 7: Results for participants' ratings of how own/other's cues reduced misunderstandings and helped anticipate the next action, visualized as boxplots with indicated median.

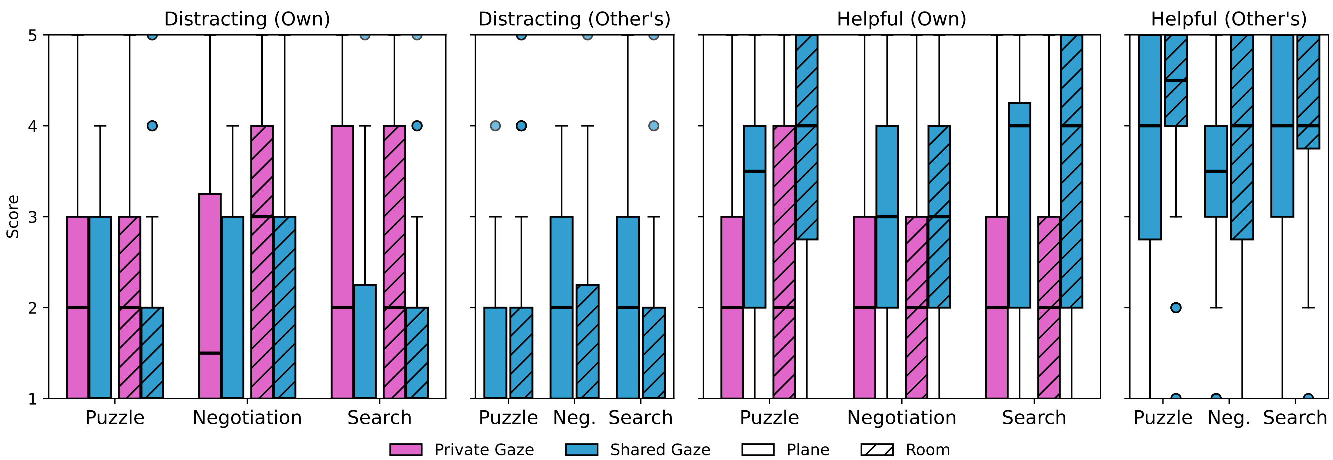


Figure 8: Results for participants' ratings of how own/other's gaze cues were distracting and helpful, visualized as boxplots with indicated median.

for a normalized measure of shared attention, see Figure 9. We analyzed the joint-attention ratio with mixed-effects models. As the outcome is bounded in (0,1), we adopted a beta mixed-effects regression with a logit link, fitted using `glmmTMB` in R. A model with random intercepts for the `group` and uncorrelated random slopes for Gaze Visibility, Interaction Space, and Task Type with an additional dispersion submodel for all IVs provided the best fit and outperformed a simpler beta model without a dispersion term. For comparison, we fitted a Gaussian LME on the raw ratio, which converged but showed singularity with richer random-effects structures. An intercept-only LME showed the same qualitative fixed-effect pattern reported below. For post-hoc comparisons, we used contrasts from the `emmeans` package. Wald tests indicated a significant main effect of Gaze Visibility ($\chi^2(1) = 13.1$, $p < .001$) with higher joint attention ratios for Shared Gaze ($M = 0.26$, $SD = 0.14$) compared to Private Gaze ($M = 0.22$, $SD = 0.13$). Further, the analysis indicated a significant main effect of Interaction Space with higher joint attention ratios for Plane ($M = 0.25$, $SD = 0.13$) compared to Room ($M = 0.22$, $SD = 0.14$). Finally, we found a third main effect of the Task Type ($\chi^2(2) = 571.3$, $p < .001$) with rising joint attention ratios from Search ($M = 0.12$, $SD = 0.06$) over Negotiation ($M = 0.22$, $SD = 0.09$) to Puzzle ($M = 0.37$, $SD = 0.10$), all $p < .001$. Moderating these main effects, we found two interactions. First, we found a Task Type × Interaction Space interaction ($\chi^2(2) = 19.2$, $p < .001$). While there was no reliable space effect for Negotiation, joint-attention ratios were higher in Plane than Room for both Puzzle ($M = 0.39$, $SD = 0.08$ vs. $M = 0.36$, $SD = 0.09$, $p < .01$) and Search ($M = 0.15$, $SD = 0.06$ vs. $M = 0.09$, $SD = 0.04$, $p < .001$). This finding is further moderated by a three-way interaction between all factors ($\chi^2(2) = 9.1$, $p < .05$).

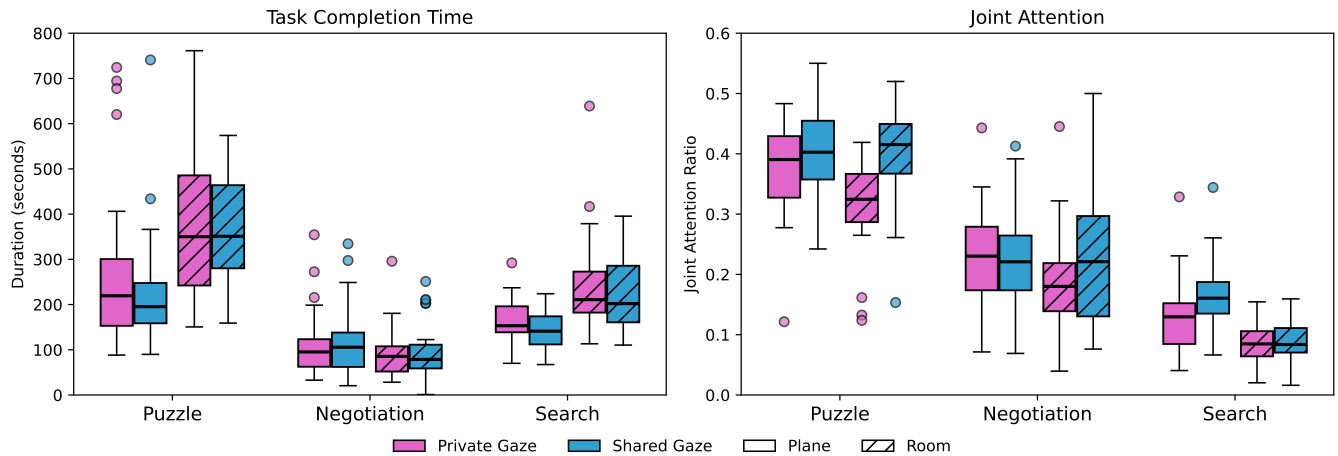


Figure 9: Results from our log data regarding Task Completion Time (left) and Joint Attention (right) visualized as boxplots with indicated median.

With Private Gaze, Plane results in rising joint-attention rates for Puzzle ($M = 0.38$, $SD = 0.08$ vs. $M = 0.31$, $SD = 0.08$, $p < .01$) and for Search ($M = 0.13$, $SD = 0.06$ vs. $M = 0.09$, $SD = 0.04$, $p < .01$). Negotiation shows the same tendency ($M = 0.23$, $SD = 0.08$ vs. $M = 0.19$, $SD = 0.09$) but does not reach significance. With Shared Gaze, differences disappear for Puzzle ($M = 0.40$, $SD = 0.07$ vs. $M = 0.40$, $SD = 0.08$, $p = 1.00$) and Negotiation ($M = 0.23$, $SD = 0.08$ vs. $M = 0.24$, $SD = 0.12$, $p = .999$), while only Search retains higher joint-attention rates for Plane ($M = 0.17$, $SD = 0.06$ vs. $M = 0.09$, $SD = 0.04$; $p < .001$).

In summary, Shared Gaze increases overall joint attention and selectively reshapes space effects, eliminating the Plane advantages for Puzzle and Negotiation across Interaction Space, but preserving a robust Plane benefit for Search.

4.4 Word Count

To investigate whether visualizing shared gaze affected how much participants spoke to each other during the tasks, we recorded participants' utterances during each task using the internal microphones of the Varjo XR-4 HMDs. We then transcribed this data, counted the words for each condition, and then normalized the data to a word per minute (WPM) rate.

We analyzed the data using a Gamma generalized linear mixed model (log link) to account for the strictly positive and right-skewed rates. The best converged model included random intercepts for `group` and for `participant` nested within `group` and fixed effects for Gaze Visibility, Interaction Space, Task Type, and their interactions. Adding by-participant slopes did not improve the fit and yielded the same substantive inferences. Wald tests indicated a single significant main effect of Task Type ($\chi^2(2) = 80.5$, $p < .001$ with WPM decreasing from Negotiation ($M = 58.70$, $SD = 23.90$) over Puzzle ($M = 45.80$, $SD = 16.50$) to Search ($M = 41.00$, $SD = 21.30$), Puzzle - Search: $p < .01$, all other $p < .001$. Neither Gaze Visibility nor Interaction Space nor any interactions were significant.

4.5 Post-Study Survey

4.5.1 Desired Future Usage. In our post-study survey, participants rated the statement: *I would like to use such a system that visualizes gaze cues of another person a lot for a similar task in the future.* for all three Task Types on an anchored 5-point Likert scale (1 = Strongly Disagree to 5 = Strongly Agree). Participants rated Negotiation lowest ($M = 2.73, SD = 1.40$), showing neither strong preference nor avoidance. In contrast, participants rated Search highest ($M = 4.29, SD = 1.03$) closely followed by Puzzle ($M = 4.12, SD = 1.16$) showing strong agreement for both Task Types.

4.6 Gaze System Quality

We also asked participants to evaluate the technical performance of the system across four dimensions: the latency and accuracy of both their own gaze visualization and that of the other participant, with ratings given on anchored 5-point Likert scales (1 = Strongly Disagree to 5 = Strongly Agree). Participants rated the acceptability of their own system latency and the others' system latency as high ($M = 4.19, SD = 0.84$, $M = 4.12, SD = 0.79$ respectively). Participants also rated accuracy positively for both one's own and other's gaze ($M = 4.04$, $SD = 0.85$, $M = 3.98$, $SD = 0.81$ respectively). These results suggest that participants were generally satisfied with both the responsiveness and accuracy of the system.

4.7 Qualitative Findings

Based on participants' comments at the end of the study regarding questions about useful and concerning situations, we looked at participants' concerns about the usage of such a system. These comments are not directly connected to our research hypotheses, but aim to extend beyond their scope, potentially allowing us to identify new directions for follow-up work. We clustered participants' feedback and synthesized it into four main concerns. **Loss of Privacy:** Participants stated a fear of being constantly observed, as gaze reveals personal information and should remain private. Participants do not want their gaze tracked or recorded without consent, and would feel exposed if anyone could see it. P153 writes: *"I don't want others to know where I am watching all the time. It is private."* Further, P149 concludes *"You can hide your thoughts less."* **Surveillance and Misuse:** Similar to above, participants were concerned that employers, teachers, or malicious actors could misuse gaze-sharing for surveillance and control. Examples include monitoring employees' focus, tracking students during exams, or extracting information (like where your valuables are or what ads draw your eye) without consent. These uses are seen as invasive and unethical. P169 writes *"Some companies could use this to fully control/check what you're doing during work hours. (...) this (...) is a invasion of privacy in my opinion."* **Distraction:** Worry that having gaze indicators on-screen or in AR could distract from primary tasks and even reduce natural communication. The extra visual stimuli might overwhelm users in complex tasks (driving, gaming), and people might rely on gaze too much and talk less, harming teamwork. *"I think it would be quite distracting with moving gaze pointers in such attentive situations."* P159 writes, referring to driving a car. On a similar notion, participants also imagined seeing gaze all the time *"could be unnecessary (...) and (...) annoying"* (P249) when visible all the time. **Social Discomfort and Misinterpretation:** Participant worried that shared gaze would lead to awkward or wrong judgments in social situations, again connecting to privacy concerns. Others might interpret one's gaze as a lack of attention, rudeness, or inappropriate interest. P165 writes *"Maybe I look at some inappropriate advertising by accident or maybe because I like it. People shouldn't know where I'm looking if I don't want them to."*. This could cause embarrassment, conflict (e.g., jealousy or offense), or the need to constantly self-censor where one looks. *"they might see that I am looking out the windows and think I am not paying attention or they might think I am looking at them seductively or at inappropriate places"* P248 worries. Participants also reflected on tradeoffs of the usage, and P173 stated *"In a short and limited task, I'd be willing to sacrifice this [privacy], but for a longer usage I don't."*

5 Discussion

We based our study design and initial assumptions on findings from related work on shared gaze in remote and collocated collaboration. Explicitly, we assume shared gaze provides efficiency gains (H1, H5) and increases the quality of collaboration (H2, H3, H4) with different impacts depending on task settings (H6).

While much of the previous work on collocated collaboration concludes performance, speed, and quality gains, upon closer investigation of the actual results, the reported data provides only limited objective measures that support these conclusions. In some instances, the comparison lacks a proper no-shared-gaze baseline, in other instances, only one of many visual implementations showed increased performance metrics while other implementations did not, making it hard to generalize to an overall performance gain through shared gaze [4, 58]. Our findings connect with the interpretation of McCarley et al. [40] of the existing body of work, stating that "*The literature in total suggests that the benefits of shared-gaze awareness to visual search may be nonrobust, or of limited generalizability.*"

5.1 Anticipation Without Acceleration

Across all three tasks and two interaction spaces of our study, shared gaze did not accelerate task completion time but did facilitate mutual anticipation in collocated AR collaboration. While prior work suggested that shared gaze can improve collaboration speed in XR, our results do not show faster task completion when a shared gaze visualization was available. Across the three task types and the two spatial arrangements we studied, completion times did not differ significantly between shared gaze and private gaze, and did not show any trends in this direction. This contradicts H1. Importantly, the gaze cue did not hinder performance either; teams were equally fast with or without shared gaze. Consistent with this, RTLX scores did not differ across visualization conditions, indicating that the gaze cue neither increased nor reduced subjective workload. However, our data only allows for a comparison between the own gaze and shared gaze condition, lacking a HMD-free non-augmented baseline to compare to. Indeed, because HMDs partially occlude the eye region and reduce visibility of eye contact and micro-cues while adding minor tracking/latency artifacts, we expect such a non-augmented, no-HMD baseline to yield even better coordination and potentially faster task completion. Several factors may explain why gaze visualization did not translate into measurable efficiency gains. Participants already coordinated effectively through speech and gestures, and task timelines were likely dominated by motor execution or individual search rather than attention alignment. The relative simplicity and non-time-critical nature of the tasks may also have limited potential benefits.

These already effective coordination channels are also reflected in our custom questionnaire items regarding the quality of communication cues, which also showed no significant differences between shared gaze and private gaze. Participants did not perceive their own or their partner's cues as more understandable, effective, intrusive, disturbing, or noticeable when shared gaze was available.

We interpret this as evidence that in collocated AR, participants rely primarily on rich natural cues (speech, gestures, and body orientation) which already provide robust grounding [8], leaving little potential for additional gains from a shared gaze. This aligns with findings from Media Richness Theory [10], which identified that face-to-face interaction is among the richest communication media, supporting immediate feedback and multiple nonverbal signals. In such settings, adding explicit gaze visualization may offer limited benefit, as the communication channel is already sufficiently rich for effective coordination. In contrast, related work in remote collaboration demonstrated robust advantages of shared gaze [17, 24]. We attribute these differences to the absence of many grounding channels in remote settings. In local settings, these cues remain available and salient, so shared gaze appears to contribute only a small amount to what is achieved through natural interaction.

However, three specific benefits in local collaboration emerged. First, shared gaze increased ratings of mutual anticipation (*CQ14: "I felt the other could anticipate my next action very well." / CQ7: "I could anticipate the others' next action very well."*). Second, participants felt their own cues reduced misunderstandings more strongly in the shared gaze condition. Third, for some tasks (Search), participants found it *CQ18: "helpful seeing the others' gaze visualized."* more than in others (Negotiation), showing that shared gaze is perceived as helpful in some instances. These effects subjectively support H2 and partially support H5. However, the effects did not translate into faster task completion (H1 not supported), indicating that gaze primarily enhances coordination quality rather than efficiency. These subjective benefits may also reflect novelty, which should be examined in longitudinal studies.

Taken together, these results imply that shared gaze in collocated AR affects quality of coordination rather than efficiency: it strengthens anticipation, reduces the perceived number of misunderstandings, and is perceived as helpful in spatial tasks, but it neither shortens completion times nor reduces workload. Collocated AR already affords effective communication through natural cues, leaving little room for gaze visualization to impact performance. The value of gaze cues, therefore, lies in subtle refinements to collaboration, not in measurable speed gains.

Gaze visualizations in collocated AR should be lightweight and complementary, supporting action anticipation and error prevention rather than replacing speech or gesture. This aligns with previous work that argues that "*understanding gaze as multimodal is critical*" [52]. When displayed in a minimal style, as in our study, they can be integrated without risking slowing down collaboration or overloading users, and may improve coordination quality in tasks with high coordination demands or spatial ambiguity. Future work should test whether these benefits persist over time, become stronger in more complex or time-critical tasks, and how they combine with other collaborative aids such as annotations or task division.

5.2 In Collocated Settings, Shared Gaze Does Not Increase Social Connectedness or Social Presence

In our study, shared gaze did not increase social connectedness beyond the Private Gaze baseline in any of the six task conditions (3 task types × 2 interaction spaces). Neither the IOS, which measures interpersonal closeness, nor the Behavioural Involvement component of the Social Presence module of the GEQ, which assesses attentiveness to a collaborator's actions, showed significant effects of gaze visualization. This contradicts H3 and provides no support for H2 on subjective measures. To the best of our knowledge, prior collocated studies with explicit gaze-sharing either assessed social connectedness only when comparing different shared-gaze visualizations without a no-shared-gaze baseline, or included such a baseline but focused on shared attention or related constructs, preventing a direct comparison.

We speculate that abundant natural cues in collocated AR (physical proximity, gesture, speech, and body orientation) already saturate the channels that gaze indicators would otherwise enhance. This result aligns with Media Naturalness Theory [30], which suggests that humans are biologically adapted for face-to-face interaction (involving speech, gesture, and facial expression) and deviations from this "natural" mode of communication increase cognitive effort and reduce social presence. While the Media Richness Theory [10] helps to explain why additional gaze cues in collocated AR may not improve performance in an already-rich medium, Media Naturalness Theory suggests that users may not experience social or cognitive gains either, since the existing interaction already matches our biologically adapted communication patterns. In our collocated AR setting, participants already shared this rich, natural channel. Adding an abstracted gaze visualization may not have increased social connection because it added no fundamentally new social cues, and possibly even competed with existing natural ones.

By contrast, in remote XR, shared gaze fills a gap by providing missing grounding signals, which explains earlier reported benefits [1, 12, 24]. Our results, therefore, highlight a boundary condition: shared gaze supports presence in socially sparse environments but becomes redundant when rich collocated cues are available.

For system design, these findings indicate that gaze visualization should not be expected to strengthen social bonds or increase presence in face-to-face AR collaboration. If enhancing social connectedness is the goal, other interventions may be more effective. At the same time, our findings do not rule out benefits in remote or hybrid settings where natural grounding cues are absent. Future work should examine whether gaze visualization yields stronger social effects when tasks are more demanding, or the environment might impede social quality.

5.3 Visualizing Shared Gaze Will Steer Attention

Participants showed significantly more joint fixations in the Shared Gaze conditions compared to Private Gaze, providing objective support for H2. This suggests that gaze visualization increased visual coordination: participants were more aware of their partner's attention and could better anticipate actions and intentions.

We interpret this as a reflexive tendency to follow gaze cues: when gaze is visualized, partners naturally look where the indicator points, producing tighter alignment of visual attention. This attentional synchrony likely improved subjective and objective markers of coordination but did not translate into faster task completion. The dissociation between coordination quality and task efficiency underscores that shared gaze changes how collaborators align attention, not how quickly they finish tasks.

From a system design perspective, this suggests gaze visualization is most useful when a tighter shared focus is critical (e.g., joint inspection, troubleshooting, or spatial search). It should not be relied upon to accelerate collaboration speed in collocated settings.

5.4 Show Us a Task Where Shared Gaze Is Objectively Useful

Participants' preferences highlight the task-dependent benefits of shared gaze. In the post-study survey, they expressed clear interest in using a hypothetical shared gaze system for future Search and Puzzle tasks, but not for Negotiation. This supports H4 and H6 and suggests that shared gaze is most valued when tasks require spatial disambiguation or joint visual search, but not when verbal reasoning strategies dominate. We speculate that participants preferred gaze visualization in Search and Puzzle because these tasks demand spatial disambiguation and deictic object reference. By contrast, Negotiation relied on dialogue and reasoning, where verbal cues were sufficient, and gaze contributed little or was even distracting. We conclude that the usefulness of gaze visualization is highly task-dependent: it enhances subjective experience in spatial and problem-solving tasks but offers little benefit in negotiation-heavy contexts. For system design, this argues against a one-size-fits-all approach. Adaptive gaze sharing should be implemented, either intelligently or with user control, so that it is available when helpful and unobtrusive when redundant.

We designed our study based on the premise that shared gaze would help both in perceived coordination and in actual efficiency, and we explicitly varied tasks and spatial arrangements to assess the strengths of these benefits. While we found evidence suggesting advances in perceived coordination, the results for the actual efficiency were not encouraging. Despite adopting and rebuilding tasks from prior work to span diverse collaboration styles in symmetric collocated settings and adding the second dimension of spatial arrangement, we found no measurable efficiency gains and not even trends for any task or space. Accordingly, we did not identify any collocated task in our set where shared gaze delivered efficiency gains. With this, we leave the community with a challenging question we were unable to answer: show us the task in symmetric collocated AR collaboration where shared gaze truly translates into objective efficiency gains. Our findings suggest that promising candidates will be tasks, where natural grounding cues are degraded or hidden because of distance or occlusion, spatial disambiguation is important for entities that are difficult to describe verbally, miscoordination impacts are strong, and speech and gesture are constrained.

In summary, our findings align with the interpretation of McCarley et al. [40], calling the current evidence "*nonrobust, or of limited generalizability*". Our findings add to this body of work, as we also could not identify any objective benefits of shared gaze.

6 Limitations and Future Work

Our study used controlled, relatively short collaborative tasks in a laboratory setting. These tasks may not capture the complexity, time pressure, interdependence, or domain specificity of many real-world AR collaborations. Future work should test longer, more complex, and time-critical task settings for more reliable results.

We conducted all trials in a room-scale environment. Findings should therefore be generalized with caution to larger, multi-room, or outdoor spaces where spatial dispersion and occlusions may increase the value of gaze visualization for maintaining shared attention. Participants interacted with physical objects in collocated AR. Whether these results transfer to fully virtual task contexts in collocated AR remains open. Virtual artifacts and avatars from hybrid settings may alter the suitability of “real-world” communication cues and the perceived role of shared gaze, potentially requiring additional communication channels.

For our baseline, we chose a private-gaze condition rather than a no-gaze condition, which allowed us to control for the effects of having a gaze cursor visible and to study the incremental effect of seeing the collaborator’s gaze. As a result, our findings describe how shared gaze adds to an interface that already includes a private gaze cursor. Future work could compare these findings to a no-gaze baseline. Further, novelty may have influenced positive attitudes toward gaze visualization. Longitudinal studies should separate initial enthusiasm from sustained utility and examine how usage patterns and perceived benefits evolve over time. Here, comfort should be assessed explicitly, given its mixed and evolving nature. Gaze sharing may feel supportive (ease, achievement) but also intrusive (perceived exposure, being watched, distraction), and these reactions may evolve over time and differ across relationships or roles. Future work should assess both positive and negative facets of comfort and privacy through longer in-the-wild studies.

We did not systematically vary or optimize combinations of cues. Future work should study interactions between gaze visualization and other modalities (annotations, deictic pointers, gesture capture, audio cues) and explore adaptive systems that enable gaze support only when beneficial. We only studied shared gaze in a dyadic setup and have no knowledge about scalability for larger groups. We speculate that in larger groups, distraction and visual clutter will be a central issue that may need novel approaches. Future work should investigate these settings and whether findings from our study transfer to these scenarios. Our two spatial setups (plane vs. room) were intended to test distinct interaction spaces, but they do not reflect the full design space. Future work should test intermediate and alternative configurations to clarify when and how interaction space changes the value of shared gaze.

7 Conclusion

We investigated the effects of shared gaze across three different task types and two interaction spaces in symmetric collocated AR collaboration. Despite this wide range of task scenarios, we did not find evidence for any benefits of shared gaze on task efficiency. The benefits of shared gaze are limited to the subjective quality of the collaboration. Our findings converge on five narratives: (1) Collocated AR already affords natural joint attention through speech and gesture, so gaze visualization adds little to social presence. (2) When displayed, gaze reliably steers attention, producing more joint fixations and stronger anticipation. (3) Shared gaze visualization does not alter perceived workload in the tasks in our study. (4) It does not reduce task completion time. (5) Nevertheless, users prefer it, especially for spatial problem-solving tasks. Taken together, these results suggest that gaze visualization improves the quality of coordination and user experience, but not efficiency – anticipation without acceleration. We therefore recommend treating shared gaze as an on-demand aid for alignment rather than an always-on accelerator for efficiency in collocated AR collaboration.

Acknowledgments

This research was supported by the Deutsche Forschungsgemeinschaft (DFG) within Project ID 529719707 "Multimodal Intent Communication of Autonomous Systems". This work has been co-funded by the LOEWE initiative (Hesse, Germany) within the emergenCITY center [LOEWE/1/12/519/03/05.001(0016)/72].

Aids

During the preparation of this work, the authors used OpenAI’s GPT-5 and Grammarly for grammar, text revision, style editing, and accessibility descriptions. For images, or parts of images, authors used OpenAI’s image generator Sora. All content was reviewed and edited by the authors, who take full responsibility for the final submission.

References

[1] Deepak Akkil, Jobin Mathew James, Poika Isokoski, and Jari Kangas. 2016. GazeTorch: Enabling Gaze Awareness in Collaborative Physical Tasks. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems*. ACM, San Jose California USA, 1151–1158. doi:10.1145/2851581.2892459

[2] Arthur Aron, Elaine N. Aron, and Danny Smollan. 1992. Inclusion of Other in the Self Scale and the structure of interpersonal closeness. *Journal of Personality and Social Psychology* 63, 4 (Oct. 1992), 596–612. doi:10.1037/0022-3514.63.4.596

[3] Huidong Bai, Prasanth Sasikumar, Jing Yang, and Mark Billinghurst. 2020. A User Study on Mixed Reality Remote Collaboration with Eye Gaze and Hand Gesture Sharing. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. ACM, Honolulu HI USA, 1–13. doi:10.1145/3313831.3376550

[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 (March 2008), 1465–1477. doi:10.1016/j.cognition.2007.05.012

[5] Frank Broz, Hagen Lehmann, Chrystopher L. Nehaniv, and Kerstin Dautenhahn. 2012. Mutual gaze, personality, and familiarity: Dual eye-tracking during conversation. In *2012 IEEE RO-MAN: The 21st IEEE International Symposium on Robot and Human Interactive Communication*. IEEE, Paris, France, 858–864. doi:10.1109/ROMAN.2012.6343859

[6] Paul-Christian Bürkner and Matti Vuorre. 2019. Ordinal Regression Models in Psychology: A Tutorial. *Advances in Methods and Practices in Psychological Science* 2, 1 (March 2019), 77–101. doi:10.1177/2515245918823199

[7] Lei Chen, Yilin Liu, Yue Li, Lingyun Yu, BoYu Gao, Maurizio Caon, Yong Yue, and Hai-Ning Liang. 2021. Effect of Visual Cues on Pointing Tasks in Co-located Augmented Reality Collaboration. In *Symposium on Spatial User Interaction*. ACM, Virtual Event USA, 1–12. doi:10.1145/3485279.3485297

[8] Herbert H. Clark and Susan E. Brennan. 1991. Grounding in communication. In *Perspectives on socially shared cognition.*, Lauren B. Resnick, John M. Levine, and Stephanie D. Teasley (Eds.). American Psychological Association, Washington, 127–149. doi:10.1037/10096-006

[9] Herbert H. Clark and Deanna Wilkes-Gibbs. 1986. Referring as a collaborative process. *Cognition* 22, 1 (Feb. 1986), 1–39. doi:10.1016/0010-0277(86)90010-7

[10] Richard L. Daft and Robert H. Lengel. 1986. Organizational Information Requirements, Media Richness and Structural Design. *Management Science* 32, 5 (May 1986), 554–571. doi:10.1287/mnsc.32.5.554

[11] Sarah D'Angelo and Andrew Begel. 2017. Improving Communication Between Pair Programmers Using Shared Gaze Awareness. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. ACM, Denver Colorado USA, 6245–6290. doi:10.1145/3025453.3025573

[12] Sarah D'Angelo and Darren Gergle. 2016. Gazed and Confused: Understanding and Designing Shared Gaze for Remote Collaboration. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. ACM, San Jose California USA, 2492–2496. doi:10.1145/2858036.2858499

[13] Sarah D'Angelo and Darren Gergle. 2018. An Eye For Design: Gaze Visualizations for Remote Collaborative Work. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM, Montreal QC Canada, 1–12. doi:10.1145/3173574.3173923

[14] S. Gareth Edwards, Lisa J. Stephenson, Mario Dalmaso, and Andrew P. Bayliss. 2015. Social orienting in gaze leading: a mechanism for shared attention. *Proceedings of the Royal Society B: Biological Sciences* 282, 1812 (Aug. 2015), 20151141. doi:10.1098/rspb.2015.1141

[15] Austin Erickson, Nahal Norouzi, Kangsoo Kim, Ryan Schubert, Jonathan Jules, Joseph J. LaViola, Gerd Bruder, and Gregory F. Welch. 2020. Sharing gaze rays for visual target identification tasks in collaborative augmented reality. *Journal on Multimodal User Interfaces* 14, 4 (Dec. 2020), 353–371. doi:10.1007/s12193-020-00330-2

[16] Ryan Khushan Ghamandi, Ravi Kiran Kattoju, Yahya Hmaiti, Mykola Maslych, Eugene Matthew Taranta, Ryan P. McMahan, and Joseph LaViola. 2024. Unlocking Understanding: An Investigation of Multimodal Communication in Virtual Reality Collaboration. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. ACM, Honolulu HI USA, 1–16. doi:10.1145/3613904.3642491

[17] Kunal Gupta, Gun A. Lee, and Mark Billinghurst. 2016. Do You See What I See? The Effect of Gaze Tracking on Task Space Remote Collaboration. *IEEE Transactions on Visualization and Computer Graphics* 22, 11 (Nov. 2016), 2413–2422. doi:10.1109/TVCG.2016.2593778

[18] Sandra G. Hart. 2006. Nasa-Task Load Index (NASA-TLX); 20 Years Later. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 50, 9 (Oct. 2006), 904–908. doi:10.1177/154193120605000909

[19] Kathryn Hays, Arturo Barrera, Lydia Ogbadu-Oladapo, Olumuyiwa Oyedare, Julia Payne, Mohotarema Rashid, Jennifer Stanley, Lisa Stocker, Christopher Lueg, Michael Twidale, and Ruth West. 2022. A state of the art and scoping review of embodied information behavior in shared, co-present extended reality experiences. *Electronic Imaging* 34, 12 (Jan. 2022), 298–1–298–19. doi:10.2352/EI.2022.34.12.ERVR-298

[20] Keita Higuch, Ryo Yonetani, and Yoichi Sato. 2016. Can Eye Help You?: Effects of Visualizing Eye Fixations on Remote Collaboration Scenarios for Physical Tasks. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. ACM, San Jose California USA, 5180–5190. doi:10.1145/2858036.2858438

[21] Wijnand A IJsselsteijn, Yvonne AW De Kort, and Karolien Poels. 2013. *The game experience questionnaire*. Technical Report. Technische Universiteit Eindhoven.

[22] Hiroshi Ishii and Minoru Kobayashi. 1992. ClearBoard: a seamless medium for shared drawing and conversation with eye contact. In *Proceedings of the SIGCHI conference on Human factors in computing systems - CHI '92*. ACM Press, Monterey, California, United States, 525–532. doi:10.1145/142750.142977

[23] Allison Jing, Kieran May, Gun Lee, and Mark Billinghurst. 2021. Eye See What You See: Exploring How Bi-Directional Augmented Reality Gaze Visualisation Influences Co-Located Symmetric Collaboration. *Frontiers in Virtual Reality* 2 (June 2021), 697367. doi:10.3389/frvir.2021.697367

[24] Allison Jing, Kieran May, Brandon Matthews, Gun Lee, and Mark Billinghurst. 2022. The Impact of Sharing Gaze Behaviours in Collaborative Mixed Reality. *Proceedings of the ACM on Human-Computer Interaction* 6, CSCW2 (Nov. 2022), 1–27. doi:10.1145/3555564

[25] Allison Jing, Kieran William May, Mahnoor Naeem, Gun Lee, and Mark Billinghurst. 2021. eyemR-Vis: Using Bi-Directional Gaze Behavioural Cues to Improve Mixed Reality Remote Collaboration. In *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems*. ACM, Yokohama Japan, 1–7. doi:10.1145/3411763.3451844

[26] Yuval Kahlon, Weiheng Hu, Momoko Nakatani, Santosh Maurya, Takuya Oki, Jiang Zhu, and Haruyuki Fujii. 2024. Immersive gaze sharing for enhancing education: An exploration of user experience and future directions. *Computers & Education: X Reality* 5 (Dec. 2024), 100081. doi:10.1016/j.cexr.2024.100081 Publisher: Elsevier BV.

[27] Jina Kang, Yiqiu Zhou, Robin Jephthah Rajarathinam, Yuanru Tan, and David Williamson Shaffer. 2024. Unveiling joint attention dynamics: Examining multimodal engagement in an immersive collaborative astronomy simulation. *Computers & Education* 213 (May 2024), 105002. doi:10.1016/j.compedu.2024.105002

[28] K. Kiyokawa, M. Billinghurst, S.E. Hayes, A. Gupta, Y. Sannohe, and H. Kato. 2002. Communication behaviors of co-located users in collaborative AR interfaces. In *Proceedings. International Symposium on Mixed and Augmented Reality*. IEEE Comput. Soc, Darmstadt, Germany, 139–148. doi:10.1109/ISMAR.2002.1115083

[29] Maurice Koch, Tobias Rau, Vladimir Mikheev, Seyda Öney, Michael Becher, Xiangyu Wang, Nelusa Pathmanathan, Patrick Gralka, Daniel Weiskopf, and Kuno Kurzhals. 2025. Group Gaze-Sharing with Projection Displays. In *Proceedings of the 2025 Symposium on Eye Tracking Research and Applications*. ACM, Tokyo Japan, 1–7. doi:10.1145/3715669.3725871

[30] N. Kock. 2005. Media Richness or Media Naturalness? The Evolution of Our Biological Communication Apparatus and Its Influence on Our Behavior Toward E-Communication Tools. *IEEE Transactions on Professional Communication* 48, 2 (June 2005), 117–130. doi:10.1109/TPC.2005.849649

[31] Ka-Wai Kwok, Loi-Wah Sun, George P. Mylonas, David R. C. James, Felipe Orihuela-Espina, and Guang-Zhong Yang. 2012. Collaborative Gaze Channelling for Improved Cooperation During Robotic Assisted Surgery. *Annals of Biomedical Engineering* 40, 10 (Oct. 2012), 2156–2167. doi:10.1007/s10439-012-0578-4

[32] Grete Helena Kütt, Kevin Lee, Ethan Hardacre, and Alexandra Papoutsaki. 2019. Eye-Write: Gaze Sharing for Collaborative Writing. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. ACM, Glasgow Scotland Uk, 1–12. doi:10.1145/3290605.3300727

[33] Gun A. Lee, Seungwon Kim, Youngho Lee, Arindam Dey, Thammathip Piumsomboon, Mitchell Norman, and Mark Billinghurst. 2017. Improving Collaboration in Augmented Video Conference using Mutually Shared Gaze. doi:10.2312/EGVE.20171359 Artwork Size: 8 pages ISBN: 9783038680383 ISSN: 1727-530X Pages: 8 pages Publication Title: ICAT-EGVE 2017 - International Conference on Artificial Reality and Telexistence and Eurographics Symposium on Virtual Environments.

[34] Jerry Li, Mia Manavalan, Sarah D'Angelo, and Darren Gergle. 2016. Designing Shared Gaze Awareness for Remote Collaboration. In *Proceedings of the 19th ACM Conference on Computer Supported Cooperative Work and Social Computing Companion*. ACM, San Francisco California USA, 325–328. doi:10.1145/2818052.2869097

[35] Yuan Li, Feiyu Lu, Wallace S Lages, and Doug Bowman. 2019. Gaze Direction Visualization Techniques for Collaborative Wide-Area Model-Free Augmented Reality. In *Symposium on Spatial User Interaction*. ACM, New Orleans LA USA, 1–11. doi:10.1145/3357251.3357583

[36] Torrin M. Liddell and John K. Kruschke. 2018. Analyzing ordinal data with metric models: What could possibly go wrong? *Journal of Experimental Social Psychology* 79 (Nov. 2018), 328–348. doi:10.1016/j.jesp.2018.08.009

[37] Roman Lissermann, Jochen Huber, Martin Schmitz, Jürgen Steimle, and Max Mühlhäuser. 2014. Permulin: mixed-focus collaboration on multi-view tabletops. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, Toronto Ontario Canada, 3191–3200. doi:10.1145/2556288.2557405

[38] Jose Llanes-Jurado, Javier Marín-Morales, Jaime Guixeres, and Mariano Alcañiz. 2020. Development and Calibration of an Eye-Tracking Fixation Identification

Algorithm for Immersive Virtual Reality. *Sensors* 20, 17 (Sept. 2020), 4956. doi:10.3390/s20174956

[39] Haiko Luepsen. 2017. The aligned rank transform and discrete variables: A warning. *Communications in Statistics - Simulation and Computation* 46, 9 (Oct. 2017), 6923–6936. doi:10.1080/03610918.2016.1217014

[40] Jason S. McCarley, Nathan Leggett, and Alison Enright. 2021. Shared Gaze Fails to Improve Team Visual Monitoring. *Human Factors: The Journal of the Human Factors and Ergonomics Society* 63, 4 (June 2021), 696–705. doi:10.1177/0018720820902347

[41] Mark B. Neider, Xin Chen, Christopher A. Dickinson, Susan E. Brennan, and Gregory J. Zelinsky. 2010. Coordinating spatial referencing using shared gaze. *Psychonomic Bulletin & Review* 17, 5 (Oct. 2010), 718–724. doi:10.3758/PBR.17.5.718

[42] 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, Amsterdam The Netherlands, 541–552. doi:10.1145/3116595.3116624

[43] Jiazhi Ou, Lui Min Oh, Jie Yang, and Susan R. Fussell. 2005. Effects of task properties, partner actions, and message content on eye gaze patterns in a collaborative task. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, Portland Oregon USA, 231–240. doi:10.1145/1054972.1055005

[44] Thammathip Piumsomboon, Arindam Dey, Barrett Ens, Gun Lee, and Mark Billinghurst. 2019. The Effects of Sharing Awareness Cues in Collaborative Mixed Reality. *Frontiers in Robotics and AI* 6 (Feb. 2019), 5. doi:10.3389/frobt.2019.00005

[45] Alexander Plopski, Teresa Hirzle, Nahal Norouzi, Long Qian, Gerd Bruder, and Tobias Langlotz. 2023. The Eye in Extended Reality: A Survey on Gaze Interaction and Eye Tracking in Head-worn Extended Reality. *Comput. Surveys* 55, 3 (March 2023), 1–39. doi:10.1145/3491207

[46] Tom Frank Reuscher, Peyman Toreini, and Alexander Maedche. 2023. The state of the art of diagnostic multiparty eye tracking in synchronous computer-mediated collaboration. *Journal of Eye Movement Research* 16, 2 (June 2023). doi:10.16910/jemr.16.2.4

[47] Rufat Rzayev, Sven Mayer, Christian Krauter, and Niels Henze. 2019. Notification in VR: The Effect of Notification Placement, Task and Environment. In *Proceedings of the Annual Symposium on Computer-Human Interaction in Play*. ACM, Barcelona Spain, 199–211. doi:10.1145/3311350.3347190

[48] Prasanth Sasikumar, Lei Gao, Huidong Bai, and Mark Billinghurst. 2019. Wearable RemoteFusion: A Mixed Reality Remote Collaboration System with Local Eye Gaze and Remote Hand Gesture Sharing. In *2019 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*. IEEE, Beijing, China, 393–394. doi:10.1109/ISMAR-Adjunct.2019.000-3

[49] Clara Sayffaerth, Annika Köhler, Julian Rasch, Albrecht Schmidt, and Florian Müller. 2025. Through the Expert's Eyes: Exploring Asynchronous Expert Perspectives and Gaze Visualizations in XR. In *2025 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. IEEE, Daejeon, Korea, Republic of, 1311–1321. doi:10.1109/ISMAR67309.2025.00136

[50] Mariah Schrum, Muyleng Ghuy, Erin Hedlund-botti, Manisha Natarajan, Michael Johnson, and Matthew Gombolay. 2023. Concerning Trends in Likert Scale Usage in Human-robot Interaction: Towards Improving Best Practices. *ACM Transactions on Human-Robot Interaction* 12, 3 (Sept. 2023), 1–32. doi:10.1145/3572784

[51] Asma Shakil, Christof Lutteroth, and Gerald Weber. 2025. A Taxonomy and Systematic Review of Gaze Interactions for 2D Displays: Promising Techniques and Opportunities. *Comput. Surveys* 57, 12 (Dec. 2025), 1–37. doi:10.1145/3736250

[52] Ludwig Sidenmark and Hans Gellersen. 2020. Eye, Head and Torso Coordination During Gaze Shifts in Virtual Reality. *ACM Transactions on Computer-Human Interaction* 27, 1 (Feb. 2020), 1–40. doi:10.1145/3361218

[53] Randy Stein and Susan E. Brennan. 2004. Another person's eye gaze as a cue in solving programming problems. In *Proceedings of the 6th international conference on Multimodal interfaces*. ACM, State College PA USA, 9–15. doi:10.1145/1027933.1027936

[54] Tianchen Sun, Yucong Ye, Issei Fujishiro, and Kwan-Liu Ma. 2019. Collaborative Visual Analysis with Multi-level Information Sharing Using a Wall-Size Display and See-Through HMDs. In *2019 IEEE Pacific Visualization Symposium (PacificVis)*. IEEE, Bangkok, Thailand, 11–20. doi:10.1109/PacificVis.2019.00010

[55] Gahyun Sung, Tianyi Feng, and Bertrand Schneider. 2021. Learners Learn More and Instructors Track Better with Real-time Gaze Sharing. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW1 (April 2021), 1–23. doi:10.1145/3449208

[56] Roel Vertegaal. 1999. The GAZE groupware system: mediating joint attention in multiparty communication and collaboration. In *Proceedings of the SIGCHI conference on Human factors in computing systems the CHI is the limit - CHI '99*. ACM Press, Pittsburgh, Pennsylvania, United States, 294–301. doi:10.1145/302979.303065

[57] Roel Vertegaal, Robert Slagter, Gerrit Van Der Veer, and Anton Nijholt. 2001. Eye gaze patterns in conversations: there is more to conversational agents than meets the eyes. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, Seattle Washington USA, 301–308. doi:10.1145/365024.365119

[58] 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 (Feb. 2017), 173–186. doi:10.1007/s00779-016-0969-x Publisher: Springer Science and Business Media LLC.

A Appendix

The open questions in the post-study questionnaire about future usage situations. Participants responded using free-text entries.

(1) Imagine you had access to a system that shows another person where you are looking in real time. In what situations in your everyday and/or work life could you see this being useful?
   (a) Please name the situation(s).
   (b) Please describe the situation(s).
   (c) Please explain why it would be useful.
(2) In what situations in your everyday and/or work life would you be concerned, hesitant, or cautious to use such a system that shows another person's gaze?
   (a) Please name the situation(s).
   (b) Please describe the situation(s).
   (c) Why would you be concerned, hesitant, or cautious?
   (d) What specific concerns would you have about using such a system?
(3) What features or controls would make you more comfortable using such a system in such situation(s)?
(4) How do you think having such gaze visualizations would change the way you interact with others in all situations you envisioned before?
(5) Compared to not having gaze visualizations, what would you gain or lose in the situations you envisioned?
(6) This is the last question. Before we conclude this study, is there anything you would like to share with us? Impressions, remarks, concerns regarding your experience in this study?

Table 3: The custom questions (CQ) after each condition.

		Strongly disagree				Strongly agree
CQ1	The other’s communication cues were very understandable.	o	o	o	o	o
CQ2	The other’s communication cues were very effective.	o	o	o	o	o
CQ3	The other’s communication cues were very intrusive.	o	o	o	o	o
CQ4	The other’s communication cues were very disturbing.	o	o	o	o	o
CQ5	The other’s communication cues were very noticeable.	o	o	o	o	o
CQ6	The other’s communication cues reduced misunderstandings a lot.	o	o	o	o	o
CQ7	I could anticipate the other’s next action very well.	o	o	o	o	o
		Strongly disagree				Strongly agree
CQ8	I felt my own communication cues were very understandable.	o	o	o	o	o
CQ9	I felt my own communication cues were very effective.	o	o	o	o	o
CQ10	I felt my own communication cues were very intrusive.	o	o	o	o	o
CQ11	I felt my own communication cues were very disturbing.	o	o	o	o	o
CQ12	I felt my own communication cues were very noticeable.	o	o	o	o	o
CQ13	I felt my own communication cues reduced misunderstandings a lot.	o	o	o	o	o
CQ14	I felt the other could anticipate my next action very well.	o	o	o	o	o
		Strongly disagree				Strongly agree
CQ15	I found it very distracting seeing my own gaze visualized.	o	o	o	o	o
CQ16	I found it very helpful seeing my own gaze visualized.	o	o	o	o	o
		Strongly disagree				Strongly agree
CQ17	I found it very distracting seeing the other’s gaze visualized.	o	o	o	o	o
CQ18	I found it very helpful seeing the other’s gaze visualized.	o	o	o	o	o