When Is Self-Gaze Helpful? Examining Uni- vs Bi-directional Gaze Visualization in Collocated AR Tasks

Daniel A. Delgado * Christopher Bowers † Rodrigo Calvo ‡ Jaime Ruiz §

Department of Computer & Information Science & Engineering
University of Florida

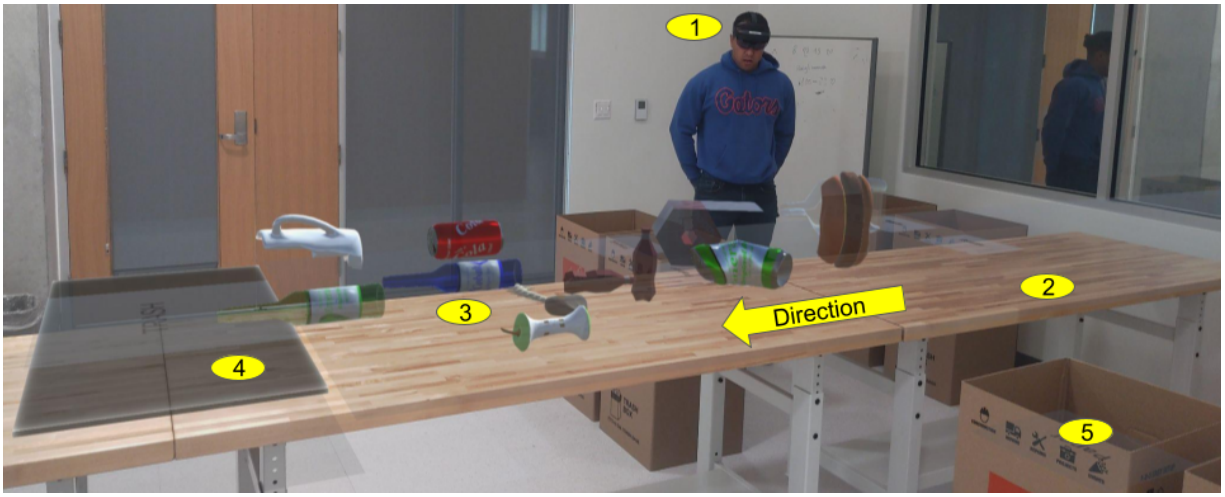


Figure 1: Overview of the physical task space with the virtual information laid over. 1) Example participant wearing the Hololens 2 Head Mounted Display. 2) Physical table that divides the participants into their respective spots and in which virtual information is laid over. 3) Virtual recycling debris that participants have to sift through and sort as it moves along the table. The debris moves along the table following the direction pointed out by the arrow. 4) As participants allow debris flow, whatever is left behind goes into the final trash bin. 5) Physical recycling bins in which the virtual debris must be sorted into.

Abstract

Shared-gaze visualizations (SGVs) in augmented reality enable collaborators to share focus and intentions through gaze interactions. Most prior research has examined bi-directional visualizations, where both users see their own and their partner's gaze, to provide feedback on how their gaze is communicated to their partner. However, bi-directional SGV approaches are largely based on research for remote collaboration. In collocated settings, bi-directional SGVs can obstruct views and cause distractions. Additionally, collocated applications differ from remote ones. We propose that if eye-tracking is well-calibrated, bi-directional visualizations may be unnecessary in collocated settings. To explore this, we conducted a user study comparing perceptions of uni- and bi-directional gaze visualizations in a virtual collaborative sorting task. Our results suggest that self-gaze may not always be necessary for users; however, there are cases in which self-gaze helps them feel more confident in the task. We offer a deeper understanding for future collaborative gaze interaction systems.

Index Terms: Shared-Gaze Visualizations, Eye-Tracking, Augmented Reality, Collocated Collaboration, Sorting

*email: danieldel@ufl.edu
†email: cj.bowers@ufl.edu
‡email: rodrigo.calvo@ufl.edu
§email: jaime.ruiz@ufl.edu

1 Introduction

Shared-gaze visualizations (SGVs) are a method for enhancing mutual understanding of collaborator focus when augmented reality headsets occlude eye contact and non-verbal communication [37]. Shared-gaze cues help collaborators predict a partner's intentions [3, 19], aiding in spatial coordination, facilitating in joint attention and aiding in collaborative awareness [1]. Gaze cues can be visualized using direct representations of eye gaze such as gaze rays [13] and cursors [21].

While existing methods have been beneficial in communicating focus, gaze cues have been implemented bi-directional for collocated applications, meaning collaborators see both their partner's and their own gaze visualization [9, 13, 21]. This approach has shown to be occasionally distracting in that it elicits more physical reactions [5] and overwhelming to users when there is too much visual information displayed [13]. In environments where space for visual content is limited, there is a need to reduce the obstruction of gaze cues while preserving their benefits. Current justifications for bi-directional visualization come from prior work in remote collaborations [21, 25]. However, there are clear apparent differences between collocated and remote collaboration. For instance, remote teams need assistance establishing trust among members due to the lack of nonverbal communication such as body language [27]. These barriers are not prevalent in collocated interactions.

One potential method for improving the distractive nature of gaze visualizations is to make only the other person's gaze visualization visible [9]. In a shared-gaze visualization interaction, visualizing a user's own gaze may be unnecessary and only the collaborator's view may need to be present in a user's field of view. Having users view their own eye gaze direction is redundant and uses valuable space in an already limited virtual environment [32].

In this work, we aim to answer whether self-gaze visualization is necessary in a joint collocated collaborative task:

1. What are the benefits/detriments of visualizing self-gaze in a collocated task?
2. Does the type of visualization affect user's perception of self-gaze?

We analyze two methods for visualizing gaze in augmented reality applications: gaze ray, a red ray extending from the user's view, and gaze hover, which highlights objects the user is looking at. Both methods are implemented in uni- and bi-directional variations.

Our work provides the following findings:

1. In an environment where a partner's gaze is already visualized, the lack of self-gaze may reduce user's own self presence, leading them to feeling left out of the task.
2. Self-gaze may not always be necessary and, in some contexts, could be more distracting than helpful.
3. Direct visualizations allow for better understanding of partner interactions and task performance, compared to more abstract or symbolic representations.

2 Related Work

Since the inception of augmented reality (AR), shared-gaze has been employed to convey collaborators' attention and focus [3]. The occlusion of faces by head mounted displays HMDs prevents users from seeing each other faces. As a means to overcome this limitation, visualization methods present to users the direction of focus from their partners. A multitude of implementations have since arisen, with more and more nuanced visualizations taking a hold. Prior to the integration of eye-tracking in HMDs, earlier research depended on head-tracking estimates to infer attention, though these methods were often imprecise [4, 29, 31]. With the introduction of eye-tracking technology in HMDs, such as the Hololens 2 [32], we can now more accurately capture users' attention and enhance visualization support. To address previous inaccuracies, gaze visualizations are provided to both collaborators and individuals themselves as feedback [21]. Shared-gaze have been mostly commonly represented through gaze rays [5, 9, 15, 16, 21, 26, 28]. Alternative methods for visualizing gaze include a 2D cursor donut [5, 21], 2D dot [20], gaze point/heatmap [1], moving trace [9], and colored highlight sphere [8].

A common theme seen from existing solutions has been the obstructive nature of gaze visualizations during tasks. For instance, gaze rays have been shown to be distracting to users [5, 25]. Additionally, users have commented on how gaze rays could be distracting in more complex tasks [13]. Alternative visualization methods like cursors [9] and highlighting objects [8, 13] still occlude a user's view.

In response, prior work has focused on designing visualization methods that complement the views of both user's themselves and their partners. For example, different visualization methods have been presented to the user compared to their partner [21], aiming to reduce distraction [37] and mental workload [25]. However, an assumption has been made based on remote collaboration studies [25], that collocated users need to be made aware of their own visualization for feedback [22]. Aside from one study in which user's gaze was visualized asynchronously [9], collocated studies of shared-gaze visualizations have mostly employed bi-directional visualizations.

Overall, uni- and bi-directional gaze visualizations have not been compared in a collocated setting. Prior work has developed gaze

(a) Bi-directional gaze ray

(b) Uni-directional gaze ray

Figure 2: Bi- and uni-directional gaze ray. In (a), the gaze ray is visualized for both ourselves and our partner. However, in (b), the gaze ray is only visualized for our partner.

visualizations using design methodologies borrowed from remote applications [25]. However, collocated settings provide different affordances that remote collaboration lacks. For instance, remote collaborators often struggle to determine where their partner is looking due to the constrained views of video cameras [24], making remote drawing annotations necessary [17].

Therefore, in this paper, we ask whether it is necessary to visualize self-gaze collocated settings. We use a series of user reported qualitative data and automatic data collection methods to understand changes in users' gaze behavior depending on uni- and bi-directional gaze visualizations. Similar to Lee et al. [25], we compare conditions in which participant's own gaze is shared and not.

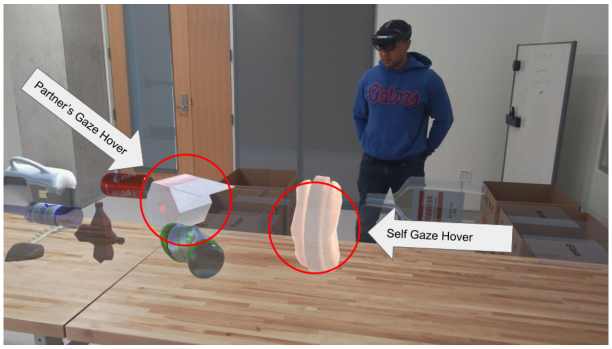
3 Methodology

We conducted a within-subjects user study in which we compared two methods of visualizing gaze through uni- and bi-directional implementations. Our study was approved by our Institutional Review Board (Protocol Number IRB202400883).

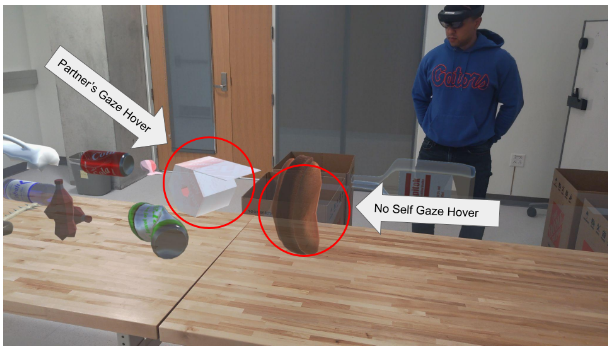
3.1 Visualizations and Directionality

The study was a 2×2 within-subjects design, varying visualization style (gaze ray vs. gaze hover) and implementation (uni-directional vs. bi-directional), resulting in four conditions. The conditions were uni-directional gaze ray, bi-directional gaze ray, uni-directional gaze hover, and bi-directional gaze hover. In uni-directional conditions, participants saw only their collaborator's visualization. In bi-directional conditions, they saw both their own and their partner's visualizations.

The gaze ray condition presents a virtual red line extending from participants head which points to their current gaze direction which mimics a similar approach used in prior work [9, 13, 28]. The gaze ray visualization is about a yard long and extends a foot out from the



(a) Bi-directional gaze hover



(b) Uni-directional gaze hover

Figure 3: Bi- and uni-directional gaze hover. In (a), the gaze hover is visualized for both ourselves and our partner, highlighting both the paper and the leftover food. In contrast, in (b), the gaze hover is visualized only for our partner, highlighting only the paper.

	Ray	Hover
Bi-directional	Condition 1	Condition 3
Uni-directional	Condition 2	Condition 4

Table 1: Experimental conditions.

users view. An example of the ray visualization is shown in Figure 2.

Our inspiration for the gaze hover came from prior implementations of gaze visualization in which objects were annotated using a bright sphere [8] and specific to the context [13]. The current gaze hover implementation provides privacy benefits in reducing what is communicated by only visualizing relevant information, instead of every eye movement [34].

In the gaze hover condition, specific objects are highlighted when users focus on them. If two users look at the same object, the object highlight stays the same. An example of each visualization is shown in Figure 3. For the bi-directional gaze condition, the participant across the table is are currently looking at the roll paper and the participant from our point of view is looking at the half-eaten burger. For the uni-directional gaze condition, the participants are looking at the same objects, however, the self-gaze is not highlighted. Therefore, the half-eaten burger remains in its original hue.

3.2 Task

The task in this study emulates a sorting process in a recycling plant [7, 35], drawing inspiration from the work of Do et al. [14]. Workers in recycling facilities could benefit from augmented reality (AR) systems that use computer vision to identify dangerous objects, thereby helping to prevent workplace injuries. Furthermore, given the typically loud environment, gaze visualizations could support non-verbal communication between workers. Similar sorting and verification processes are also common in various other industrial applications [12, 33].

This study simulates a scenario in which AR headsets equipped with computer vision tools are used to detect hazardous objects in a recycling plant sorting facility. Workers' communication is supported by shared-gaze visualizations within the AR headsets to aid non-verbal signaling.

Figure 4: The dangerous objects—such as broken glass bottles or metal cans—are highlighted among the debris of the rest of the recycling material.

The task involves users sorting virtual objects as they float down a 3D stream over a physical table. The goal is to sort each object into the appropriate bin based on its designated shape—otherwise, users lose points. Undesirable or dangerous materials must be left untouched. These dangerous materials are highlighted with a red hue 80% of the time to emulate current computer vision detection accuracies [30, 36] (Shown in Figure 4). A layout of the room is presented in Figure 5.

The objects in the scene were set up to move at a speed of 0.3125 meters per second in Unity. The position of the objects was adjusted at 16 millisecond intervals. Each adjustment altered the position by 0.005 meters, thus giving a theoretical speed of 0.3125 m/s. However, the actual observed speed of the objects resulted in being closer to 0.2 m/s in practice. Differences between the theoretical and actual speeds can be attributed to latency within the system used.

Participants are asked to sort four types of recyclable materials: glass, metal, paper, and plastic. An example of the materials is illustrated in Figure 1. Additionally, participants must avoid undesirable items (e.g., broken glass and uneaten food). Participants are free to use shared-gaze visualizations and to communicate using any modality they find comfortable.

Before each task, participants are briefed on the nature of the task and the visualization style being used. They are also encouraged to coordinate their actions. In the virtual environment, only one user can interact with an object at a time. As a result, participants must collaborate—if both attempt to grab the same object, it becomes suspended until one user releases it. This constraint promotes the use of shared-gaze visualizations and other available communication modalities.

Scoring is shared between participants, and success is evaluated collectively rather than individually. The virtual environment uses Unity colliders to detect interactions within the task—for example, when one object touches another. Each time an object is correctly sorted into its corresponding bin, a Unity collider logs the event to a file.

3.3 Procedure

Participants were recruited from a local university campus through flyers, email recruitment, and word of mouth. Participants could sign up either as a pair or individually. If participants were unable to present as a pair, they were provided with a confederate, supported by a researcher [10]. Out of the 21 participants recruited, only three performed the study with a confederate. Only data from the recruited participants was used for analysis.

The study began by obtaining informed consent from participants. We then handed each participant a Hololens 2 and performed the device's built-in eye calibration [2]. Once calibration was complete, a researcher would begin our study application. We then performed a pre-study test to ensure the physical and virtual objects were aligned, eye-tracking was functioning, and participants understood the task. The study began once participants could pick up virtual objects using pinch gestures and correctly identify which objects should be recycled.

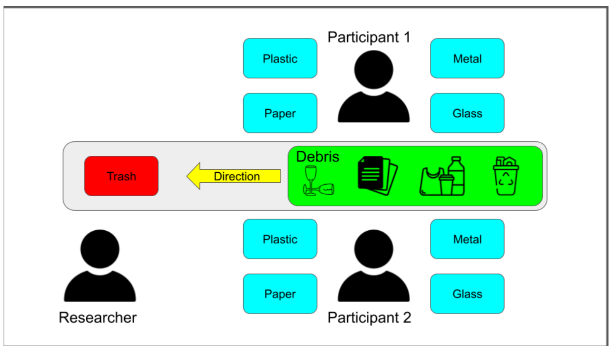


Figure 5: Top-down diagram of the task layout which includes bins for recycling material, location of participants within the task, and the direction of debris flow.

Participants were given five minutes to sort as many virtual recyclable items as they could. After the five minutes, participants were given a collaborative feedback survey. This process was then repeated for each condition. Conditions were counterbalanced across participants using a Balanced-Latin Square. At the end of the four conditions, participants were given the visualization comparison and feedback, and demographics surveys. The study took approximately an hour and thirty minutes to complete.

3.4 Data Collection

For our study, we used both quantitative and qualitative measures. For quantitative measures, we recorded task performance, measured by the number of objects correctly sorted in each condition. We also tracked how often participants interacted with a dangerous object. The Unity application logged each instance of these interactions in a file.

For qualitative measures, we collected a series of collaborative feedback, visualization comparison and feedback, and demographic surveys. The application collected user interactions throughout the study which included the number of objects sorted and the number of dangerous objects interacted with. For the collaborative feedback survey, we used a modified version of a collaborator workload survey from prior work [13, 21] (Shown in Table 2). Participants were asked to rate their agreement with a set of statements on a scale from 1 to 100. An open-ended response section was included after the agreement statements to allow participants to explain their ratings. Additionally, we collected a visualization comparison and feedback survey which has participants rank the shared-gaze visualizations and provide a reasoning for their choices. Finally, we collected a demographics survey that collected information on participants' experiences with mixed reality and sorting tasks from current or prior jobs in their personal lives.

(a) Paper and plastic bins

(b) Metal and glass bins

Figure 6: Physical recycling bins overlaid with virtual information

Question No.	Survey Element
Intentions	
Q1	"My intentions are accurately represented"
Q2	"My partners intentions are accurately represented to me"
Q3	"I can understand my partners' focus with ease"
Focus	
Q4	"It is better for me to understand my partner's focus"
Q5	"It is better for my partner to understand my focus"
Attention	
Q6	"It is easy to observe my partner's attention"
Q7	"It is easy for my partner to observe my attention"
Reaction	
Q8	"I react to my partner frequently"
Q9	"My partner reacts to me frequently"
Interaction	
Q10	"This form of visualization is effective"
Q11	"This form of visualization is engaging"

Table 2: Collaborative feedback survey questions.

3.5 Apparatus

The system was implemented using Unity Engine 2020.3.20f1 running on a Lenovo Yoga Laptop. All command controls were made using the laptop during the study. Two Hololens 2 were used for the main augmented reality platform [32]. The Hololens 2 has built in eye-tracking and the appropriate calibration [2]. For markerless interaction in the physical world, Microsoft Azure Spatial Anchors were used for syncing the environments between headsets [6]. For networking between devices, Photon Engine [11], was used to communicate asynchronously.

3.6 Participants

We recruited 21 participants (14 Male, 7 Female) between the ages of 18 and 41 years old (mean 25.43 years, SD = 5.38 years). Most of the participants had some prior experience with a mixed reality system beforehand (14 Yes, 7 No). A majority of participants knew their study partner (14 Yes, 7 No). Additionally, a majority of participants had no experience with sorting tasks (8 Yes, 13 No). Participation in our study was completely voluntary. However, participants were allowed to receive extra credit for applicable courses.

4 Results

In this section we present quantitative and qualitative results from our surveys. We compared our data by order of directionality (uni vs bi), visualization style (gaze vs hover), and the interaction effects between the two.

4.1 Task Performance

A two-way repeated measures ANOVA was conducted to examine the effect of *Visualization* and *Directionality* on the number of objects sorted and the interaction effects between visualization and directionality. The analysis revealed a significant main effect of Visualization ($F_{(1,11)} = 5.768, p < 0.05$), but no significant effect of Directionality ($F_{(1,11)} = 0.164, p = 0.693$) or their interaction ($F_{(1,11)} = 0.051, p = 0.826$).

Visualization	Directionality	Mean	SD	N
Ray	Uni	34.7	18.2	12
Ray	Bi	33.1	15.6	12
Hover	Uni	25.4	10.4	12
Hover	Bi	25.1	10.4	12

Table 3: Descriptive statistics for the number of objects sorted across conditions.

Factor	$F(1,11)$	p-value	η_g^2
Visualization	5.768	**0.035**	0.093
Directionality	0.164	0.693	0.001
Interaction (Vis. × Dir.)	0.051	0.826	0.0005

Table 4: ANOVA results for Visualization, Directionality, and their interaction on the number of objects sorted.

Visualization (Ray vs. Hover) There is a statistically significant effect of visualization style on the number of sorted objects ($F_{(1,11)} = 5.77, p = 0.035, ges = 0.093$). The effect size (generalized eta-squared, ges = 0.093) indicates a small-to-moderate effect. This suggests that participants sorted a different number of objects depending on whether they used the Ray or Hover visualization.
Directionality (Uni vs. Bi-directional) The effect of directionality is not significant. This means there is no strong evidence that sorting performance was affected by whether the visualization was uni- or bi-directional.
Interaction (Visualization × Directionality) The interaction effect is not significant. This suggests that the effect of visualization does not depend on whether the condition was uni- or bi-directional. The effect size is very small, meaning there is little practical impact of this interaction.

Descriptive statistics suggest minor differences between conditions in the number of dangerous objects participants interacted with, though not statistically significant. The mean number of interactions per condition were:

Visualization	Directionality	Mean	SD	N
Ray	Uni	1.08	1.88	12
Ray	Bi	1.08	1.73	12
Hover	Uni	1.33	2.27	12
Hover	Bi	0.75	0.97	12

Table 5: Descriptive statistics for interactions with dangerous objects across conditions.

A two-way repeated measures ANOVA was conducted to examine the effect of Visualization and Directionality on the number of times users interacted with dangerous objects. The results suggest that neither Visualization nor Directionality had a significant effect on the number of interactions with dangerous objects. High standard deviations indicate substantial variability among participants.

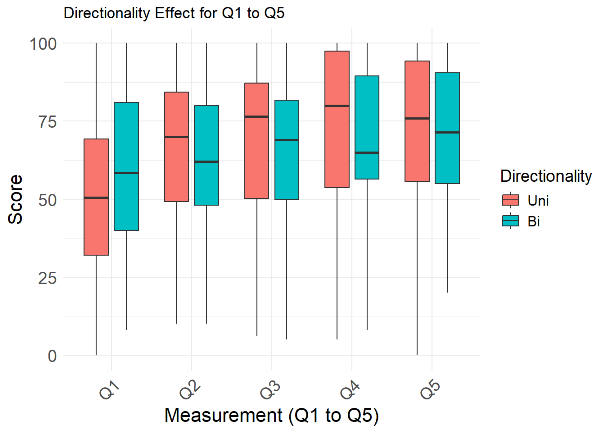
4.2 Reliability Scale of Collaborative Feedback Survey

To assess the internal consistency of the 11-item collaborative feedback questionnaire (Q1 to Q11), we computed Cronbach's alpha. The scale demonstrated excellent reliability, with a raw Cronbach's alpha of 0.94. The standardized alpha was nearly identical ($\alpha = 0.94$), and Guttman's Lambda 6 further supported the scale's reliability ($\lambda_6 = 0.97$). The average inter-item correlation was 0.60, with a median inter-item correlation of 0.55, indicating strong item coherence.

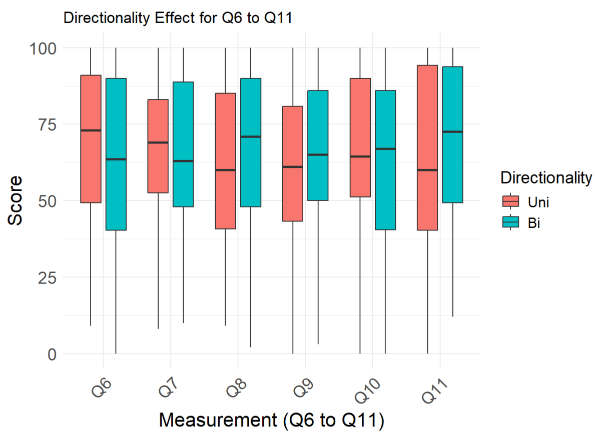
4.3 Analysis of Collaborative Feedback Survey Responses

The results from the collaborative experience survey did not meet the parametric assumptions required for ANOVA. Therefore, we applied an ArtANOVA test [23] to compare the effects of directionality and visualization style, treating the data as paired within subjects. Significant effects are reported based on p-values. Results from the survey are shown in Figures 7a and 7b for directionality, and Figures 8a and 8b for visualization.

The ArtANOVA analysis examined the effects of Directionality, Visualization, and their Interaction across the eleven collaborative feedback survey measures (Q1 to Q11) in Table 2. The results are summarized as follows:
Visualization (Ray vs. Hover) Visualization showed significant effects in 7 out of 11 measures. The strongest effects were observed in Q2 "My partners intentions are accurately represented to me" ($F_{(1,60)} = 12.91, p = 0.001$), Q3 "I can understand my partners' focus with ease" ($F_{(1,60)} = 12.15, p = 0.001$), Q6 "It is easy to observe my partner's attention" ($F_{(1,60)} = 7.57, p = 0.008$), and Q10 "This form of visualization is effective" ($F_{(1,60)} = 8.26, p = 0.006$), all of which reached a high level of statistical significance ($p < 0.01$). Additional significant effects ($p < 0.05$) were found in Q1 "My intentions are accurately represented" ($F_{(1,60)} = 6.61, p = 0.013$), Q4 "It is better for me to understand my partner's focus" ($F_{(1,60)} = 5.51, p = 0.022$), Q5 "It is better for my partner to understand my focus" ($F_{(1,60)} = 4.96, p = 0.030$), and Q7 "It is easy for my partner to observe my attention" ($F_{(1,60)} = 6.43, p = 0.014$). In contrast, Q8 ($F_{(1,60)} = 2.25$), Q9 ($F_{(1,60)} = 1.26$), and Q11 ($F_{(1,60)} = 3.40$) did not reach significance. These results suggest that the visualization style had a strong influence on participant responses, particularly in Q2, Q3, Q6, and Q10, where effects were highly significant.

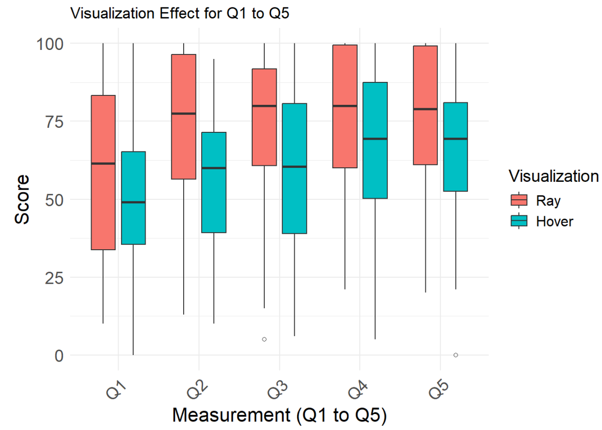


(a) Survey responses Q1 to Q5.

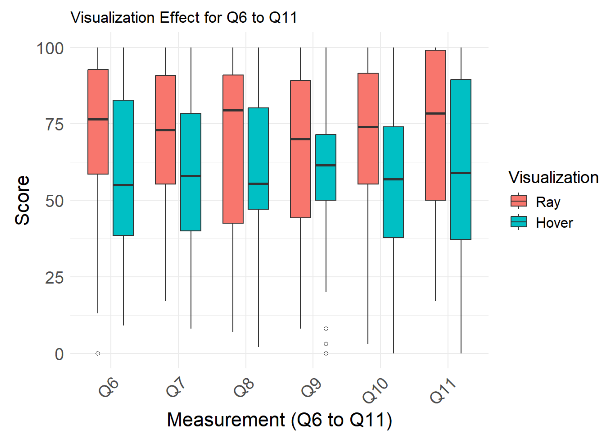


(b) Survey responses Q6 to Q11.

Figure 7: Collaborative feedback survey responses for directionality. A rating of 0 meant participants did not agree with the statement at all, while a rating of 100 meant they fully agreed.



(a) Survey responses Q1 to Q5.



(b) Survey responses Q6 to Q11.

Figure 8: Collaborative feedback survey responses for visualization. A rating of 0 meant participants did not agree with the statement at all, while a rating of 100 meant they fully agreed.

The results indicate that the visualization style (Ray vs. Hover) significantly influenced participants' perceptions in multiple areas related to mutual awareness and effectiveness. The strongest effects ($p < 0.01$) were found in measures related to understanding a partner's ATTENTION (Q6) and INTENTION (Q2, Q3) and INTERACTION (Q10). This suggests that Ray provided a clearer representation of a partner's gaze and attention, improving participants' ability to interpret their partner's focus with ease.

Additional significant effects ($p < 0.05$) in Q1, Q4, Q5, and Q7 further reinforce this trend, showing that visualization impacted how well participants felt their own focus was conveyed and how effectively they could interpret their partner's gaze. However, measures related to broader observational aspects (Q8, Q9, Q11) did not reach significance, indicating that while visualization strongly influenced direct perception and ease of understanding, it may have had less impact on other aspects of the interaction.

Overall, these results highlight the importance of visualization style in facilitating gaze awareness and mutual understanding, particularly in direct measures of focus comprehension.

Directionality (Uni vs. Bi-directional) Directionality did not reach statistical significance in any measure ($p > 0.05$ for all tests).

Interaction (Visualization × Directionality) Interaction effects were non-significant across all measures, with $p > 0.05$ in every case. These findings indicate that the effect of Visualization was independent of Directionality, meaning that the visualization style had a consistent effect across conditions regardless of whether gaze directionality was uni- or bi-directional.

Key takeaways from the analysis are as follows. First, Visualization had a significant effect on many of the collaborative feedback survey measures, particularly Q2, Q3, Q6, and Q10 (Table 2). Second, Directionality did not significantly impact responses. Finally, no significant interaction effects were observed, meaning that Directionality did not modify the impact of Visualization. These findings suggest that the type of visualization influenced participants' responses, but whether gaze directionality was uni- or bi-directional did not play a major role, with Q1 being the only measure almost reaching significance. Additionally, interaction effects between both directionality and visualization were not significant. Results are summarized in Table 6.

4.4 Qualitative Feedback

Each collaborative feedback survey contained open response section which allowed participants to describe their choices. These responses give us a deeper insight into participant's experiences throughout the collaborative interactions.

Although not significant, results from the collaborative feedback survey suggested participants felt bi-directionality more accurately represented their intentions compared to uni-directional. A look into the open responses sheds some light as to participants' perceptions.

Throughout the study, participants used gaze visualizations to

Measure	Effect	*F*	*p*	Significance
Q1	Directionality	3.12	0.082	.
	Visualization	6.61	**0.013**	*
	Interaction	0.13	0.718	
Q2	Directionality	1.00	0.322	
	Visualization	12.91	**0.001**	***
	Interaction	0.10	0.753	
Q3	Directionality	0.71	0.403	
	Visualization	12.15	**0.001**	***
	Interaction	0.01	0.907	
Q4	Directionality	2.42	0.125	
	Visualization	5.51	**0.022**	*
	Interaction	0.04	0.834	
Q5	Directionality	0.18	0.675	
	Visualization	4.96	**0.030**	*
	Interaction	0.01	0.919	
Q6	Directionality	0.83	0.367	
	Visualization	7.57	**0.008**	**
	Interaction	0.02	0.882	
Q7	Directionality	0.78	0.380	
	Visualization	6.43	**0.014**	*
	Interaction	1.00	0.322	
Q8	Directionality	0.96	0.331	
	Visualization	2.25	0.139	
	Interaction	0.58	0.449	
Q9	Directionality	1.01	0.319	
	Visualization	1.26	0.266	
	Interaction	3.36	0.072	.
Q10	Directionality	0.41	0.524	
	Visualization	8.26	**0.006**	**
	Interaction	0.29	0.590	
Q11	Directionality	0.18	0.671	
	Visualization	3.40	0.070	.
	Interaction	1.37	0.247	

Table 6: Summary of ArtANOVA Results for the collaborative feedback survey.

signal their intentions to grab specific objects, particularly relying on the uni-directional gaze hover to avoid overlap with their partner. However, the lack of self-view in gaze hover reduced confidence, as P9 commented, ”I did not like that it was not highlighted ... made me feel like I was doing nothing at all ...” Participant 2 shared similar sentiments remarking that gaze ray in the bi-directional condition was the ”Best of the alternatives since it clearly showed my attention and my partner’s while making them distinct to each other. Additionally, the feedback of what I was looking at in addition to the grabbing motion/feedback of the HoloLens which make me more aware of my own attention and purpose.”

We see a similar interested for self-gaze gaze in the hover conditions as well with participant 21 stating:

> ”I like this shared-gaze visualization the most. The line visualization cluttered the scene. This highlighting visualization technique was a bit confusing because I wasn’t sure if things I was picking up were highlighted or things I was looking at were highlighted. Also I feel that gaze detection is not as accurate as detecting things I’m picking up. Still, this technique allowed me and my collaborator to not need to communicate, we got through all the items quickly and focused on recycling items that were closer to us and the highlights helped us know what doesn’t need to be picked up because the collaborator is handling it.”

Although results indicated that participants favored the bi-directional conditions, we still observed notable contradictions to bi-directionality. For instance, participants found the self-gaze to be unfavorable describing the bi-directional condition to be ”Very useful since it lets me see where they’re looking at, [however] slightly confusing since I see my gaze too.” Participants additionally enjoyed the reduced virtual content in uni-directional conditions, with participant 12 noting the uni-directional gaze ray ”...effective at showing where my partner was looking, without the distraction of my own focus.” We noticed participants feeling like the self-gaze was distracting with participant 12 also remarking the uni-directional gaze-hover condition ”... was very effective at showing where my partner was looking, without the distraction of my own focus.”

In regards to visualization, participants rated the ray condition higher than the hover condition for measures Q1-Q7 and Q10. Participants felt the gaze ray represented their intentions and their partners’ more accurately, represented their partner’s focus with ease, allowed for them to understand their partner’s focus and vice verse, and made it easier to observe their partner’s attention and vice versa. Finally, the gaze ray condition was rated to be a more effective form of visualization compared to gaze hover.

A look into open responses highlights participants opinions. Participants found it easier to grasp their partner’s focus with gaze ray conditions, especially in the uni-directional case. For example, one participant stated, ”I think this one [uni-directional gaze] is better than the others [conditions] because it lets me see what he is seeing without overloading my view. I don’t really need to know what I’m looking at, but what he’s looking at helps me (P20).”

The absence of cues associating each gaze hover led to confusion about whose visualization was whose, requiring participants to take extra steps to communicate their actions. For instance, Participant 3 remarked, the gaze hover visualization was ”Very ineffective since I had to communicate using my voice more with my partner.”

4.5 Post Study Survey

In the post study survey, we asked participants to rank their preference of visualization type. Based on these results, we ran a Wilcoxon Signed-Rank Test and found no significant difference between preferences of the uni- and bi-directional gaze ray and hover conditions. In regards to understanding participants view on the visual cues, we found that a slight majority of them found them helpful for identifying dangerous objects in context of the shared-gaze visualizations. Participants gave mixed opinions on their perceptions of the visual cues . For instance, participant 13 noted how the cues ”... made the objects more detectable, and therefore avoidable”. However, participant 16 stated ”[I] didn’t really notice them.” Overall, participants were more focused on the core task and the shared-gaze visualizations.

4.6 Summary of results

When considering task performance, the type of visualization had a significant effect on the number of objects participants sorted, with participants sorting more objects when using the ray visualization compared to the hover visualization. Directionality had a slight effect on measure Q1 from the collaborative feedback survey, suggesting that participants felt their actions were better represented when using bi-directional visualizations compared to uni-directional ones; however, this result was not statistically significant. The type of visualization had a significant effect on measures Q1–Q7 and Q10 from the collaborative feedback survey. Overall, these findings highlight the impact of visualization style on gaze awareness and mutual understanding, particularly in the comprehension of focus. Open-ended responses from the collaborative feedback survey provided deeper insights into participant interaction.

Choice	Ranked 1st	Ranked 2nd	Total Responses
Ray	14	7	21
Hover	7	14	21

(a) Survey results for "Rank the shared-gaze visualization technique."

Choice	Ranked 1st	Ranked 2nd	Total Responses
Ray	12	9	21
Hover	9	12	21

(b) Survey results for "Rank the shared-gaze visualization technique with in the context of sorting."

Response	Count
Yes	12
No	9

(c) Survey responses for "Were the visual cues helpful in avoiding dangerous items?"

Table 7: Results for the visualization comparison and feedback survey.

5 Discussion

Self-gaze in a collocated task Results showed no significant difference between measures when comparing them by directionality. Measure Q1 in the collaborative feedback survey showed a trend with participants suggesting that their intentions were more accurately represented while using the bi-directional conditions. Participants provided valid reasons for using and opting-out of self-gaze in collaborative interactions. A leading reason for participants preferring self-gaze was otherwise feeling left out and reducing their confidence. This is consistent with prior work in remote work in which participants felt greater presence within the task when their own visualization was visible [18, 25].

Visualization and directionality Compared to prior work in remote gaze visualization [25], we compare two different methods of self-gaze. While participants were able to sort more virtual objects while using the ray visualization compared to the hover visualization, overall quantitative results suggest that the type of visualization did not affect user perceptions of directionality. Interaction effects between directionality and visualization were not significant.

5.1 Showing Ownership of Gaze Visualizations

The confusion associated with gaze hover may stem from participants' unfamiliarity with whose visualization represented whom. The uni-directional conditions aided understanding, as participants knew only their partner's view would be visualized. In the bi-directional conditions, however, participants often needed to take extra steps to clarify their intentions and actions to avoid conflicting with their partner—for instance, reaching for the same object due to mistaken assumptions about ownership of the visual cue. The difference in ease between gaze ray and gaze hover can be attributed to participants' ability to associate the ray with their partner, as it originates from the other person's head and reorients with their movements. When designing a gaze visualization where ownership is not immediately clear, a uni-directional approach may help. Alternatively, adding an external cue that explicitly shows ownership can help reduce confusion. For example, prior implementations have used color to indicate ownership [21]. However, adding visual elements may risk increasing distraction rather than aiding clarity.

5.2 Combination of Techniques

One approach to overcome the confusion of discerning between one's own visualization and their partner's would be to use a combination of visualizations for users and their partners. In this study, we observed shared-gaze visualizations in a context in which both the user's and their partner's visualizations were the same. By implementing different visualization techniques for each collaborator, we could preserve the user's sense of presence while also minimizing distracting elements and confusion about visualization ownership. For example, we could visualize a ray for our partners to provide a direct and clear indication of their intentions, while using the hover visualization for ourselves, which provides a more subtle reference.

5.3 On/Off Gaze Visualizations

A compromise to uni- and bi-directional gaze visualization would be turning on self-gaze at necessary points during an interaction. By allowing participants to choose when they want to see would provide them the benefits associated with self-gaze, such as confidence from feedback [25] and reassurance of fair participation as seen from our results, while removing the distractions of constant view of self-gaze. Methods for controlling the current state of gaze visualizations could be implemented through multi-modal interactions, however it is important to consider the importance of each modality to natural human interactions. For instance, voice control could interfere with verbal communication, while gestural control could restrict the user's body movements. If designing automatic methods of visualizing gaze visualizations, it is important to consider the frequency of activation. More clutter in an already dynamic environment can prove to be more distracting than beneficial.

6 Limitations and Future Work

Our study provided insights into users' preferences and understanding of unidirectional versus bidirectional gaze visualizations, but several limitations should be noted.

The HoloLens 2 devices we used introduced technical constraints, including limited rendering capacity for complex tasks, increasing latency over time, and low display brightness that made some elements appear semi-transparent. Future work should consider alternative hardware better suited for rendering dense or prolonged SGV interactions.

While our SGV implementations (gaze ray and hover) represent distinct approaches, they do not capture the full range of possible visualizations. Future research should evaluate additional methods, such as gaze outlines or trigger-based cues, to broaden understanding of user preferences.

Moreover, our study focused on a single collaborative task. Although participants used multimodal cues—such as gestures and vocal communication—along with SGVs, the scope of application contexts was narrow. Future studies should explore SGV use across a wider range of tasks, particularly those involving physical objects, to examine how SGVs support real-world applications. Additionally, our study observed only one instance of object speed throughout all conditions. Future work could investigate varying speeds to observe how increased intensity of the tasks affects performance through stress-testing.

The conditions in our study looked at settings in which a shared-gaze visualization was always present. Future research should consider investigating how a condition with the absence of a visualization could affect users' perceptions throughout the study.

Finally, while we did not include a formal behavioral analysis, our findings are supported by participant comments. Future work could extend these results by conducting a behavioral analysis of video data.

7 Conclusion

In this paper, we present the findings of a user evaluation of uni- and bi-directional gaze ray and gaze hover during a collaborative industrial task in augmented reality. Our results showed that participants occasionally found it more convenient to view only their partner's shared-gaze visualization while completing a dynamic task. However, the lack of self-gaze visualization presents some hurdles when their partner's visualization is visible. For example, participants felt like they were not participating in the task and relied on self-gaze as a feedback cue to reassure themselves of what their partner were seeing. Through these findings, we provide some design considerations for implementing uni- and bi-directional gaze ray and hover.

Acknowledgments

This work is partially supported by the National Science Foundation Award #IIS-1750840. Any opinions, findings, and conclusions or recommendations expressed in this paper are those of the authors and do not necessarily reflect these agencies' views.

References

[1] D. Akkil and P. Isokoski. Gaze augmentation in egocentric video improves awareness of intention. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, pp. 1573–1584, 2016. 1, 2

[2] S. Aziz and O. Komogortsev. An assessment of the eye tracking signal quality captured in the hololens 2. In *2022 Symposium on eye tracking research and applications*, pp. 1–6, 2022. 4, 5

[3] M. Billinghurst, S. Weghorst, and T. Furness. Shared space: An augmented reality approach for computer supported collaborative work. *Virtual Reality*, 3:25–36, 1998. 1, 2

[4] N. Binetti, T. Cheng, I. Mareschal, D. Brumby, S. Julier, and N. Bianchi-Berthouze. Assumptions about the positioning of virtual stimuli affect gaze direction estimates during augmented reality based interactions. *Scientific Reports*, 9(1):2566, 2019. 2

[5] L. Brägger, L. Baumgartner, K. Koebel, J. Scheidegger, and A. Çöltekin. Interaction and visualization design considerations for gaze-guided communication in collaborative extended reality. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 4:205–212, 2022. 1, 2

[6] A. Brunzini, M. Ciccarelli, M. Sartini, G. Menchi, A. Papetti, and M. Germani. A novel approach to use marker-less mixed reality applications with in-motion systems. In *International Joint Conference on Mechanics, Design Engineering & Advanced Manufacturing*, pp. 1401–1412. Springer, 2022. 5

[7] Centers for Disease Control and Prevention. Wholesale trade recycling workers: Injuries and prevention. `https://blogs.cdc.gov/niosh-science-blog/2020/07/01/wholesale-recycling/`, 2020. Accessed: 2024-08-21. 3

[8] Y. Cha, S. Nam, M. Y. Yi, J. Jeong, and W. Woo. Augmented collaboration in shared space design with shared attention and manipulation. In *Adjunct Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology*, pp. 13–15, 2018. 2, 3

[9] L. Chen, Y. Liu, Y. Li, L. Yu, B. Gao, M. Caon, Y. Yue, and H.-N. Liang. Effect of visual cues on pointing tasks in co-located augmented reality collaboration. In *Proceedings of the 2021 ACM Symposium on Spatial User Interaction*, pp. 1–12, 2021. 1, 2

[10] L.-P. Cheng, P. Lühne, P. Lopes, C. Sterz, and P. Baudisch. Haptic turk: a motion platform based on people. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 3463–3472, 2014. 4

[11] Q. Dao. Multiplayer game development with unity and photon pun: a case study: Magic maze. 2021. 5

[12] A. G. De Sa and G. Zachmann. Virtual reality as a tool for verification of assembly and maintenance processes. *Computers & Graphics*, 23(3):389–403, 1999. 3

[13] D. A. Delgado and J. Ruiz. *Evaluation of Shared-Gaze Visualizations for Virtual Assembly Tasks*. IEEE VR, 2024. 1, 2, 3, 4

[14] T. D. Do, S. Y. Dylan, A. Katz, and R. P. McMahan. Virtual reality training for proper recycling behaviors. In *ICAT-EGVE (Posters and Demos)*, pp. 31–32, 2020. 3

[15] A. Erickson, N. Norouzi, K. Kim, J. J. LaViola, G. Bruder, and G. F. Welch. Effects of depth information on visual target identification task performance in shared gaze environments. *IEEE transactions on visualization and computer graphics*, 26(5):1934–1944, 2020. 2

[16] A. Erickson, N. Norouzi, K. Kim, R. Schubert, J. Jules, J. J. LaViola Jr, G. Bruder, and G. F. Welch. Sharing gaze rays for visual target identification tasks in collaborative augmented reality. *Journal on Multimodal User Interfaces*, 14(4):353–371, 2020. 2

[17] S. Gauglitz, B. Nuernberger, M. Turk, and T. Höllerer. In touch with the remote world: Remote collaboration with augmented reality drawings and virtual navigation. In *Proceedings of the 20th ACM Symposium on Virtual Reality Software and Technology*, pp. 197–205, 2014. 2

[18] K. Gupta, G. A. Lee, and M. Billinghurst. 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):2413–2422, 2016. 8

[19] K. Higuch, R. Yonetani, and Y. Sato. 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*, pp. 5180–5190, 2016. 1

[20] P. Jansen, F. Fischbach, J. Gugenheimer, E. Stemasov, J. Frommel, and E. Rukzio. Share: Enabling co-located asymmetric multi-user interaction for augmented reality head-mounted displays. In *Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology*, pp. 459–471, 2020. 2

[21] A. Jing, K. May, G. Lee, and M. Billinghurst. Eye see what you see: Exploring how bi-directional augmented reality gaze visualisation influences co-located symmetric collaboration. *Frontiers in Virtual Reality*, 2:697367, 2021. 1, 2, 4, 8

[22] A. Jing, K. May, B. Matthews, G. Lee, and M. Billinghurst. The impact of sharing gaze behaviours in collaborative mixed reality. *Proceedings of the ACM on Human-Computer Interaction*, 6(CSCW2):1–27, 2022. 2

[23] M. Kay and J. O. Wobbrock. Package 'artool'. *CRAN Repository*, 2016:1–13, 2016. 5

[24] S. Kim, G. A. Lee, and N. Sakata. Comparing pointing and drawing for remote collaboration. In *2013 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 1–6. IEEE, 2013. 2

[25] G. A. Lee, S. Kim, Y. Lee, A. Dey, T. Piumsomboon, M. Norman, and M. Billinghurst. Improving collaboration in augmented video conference using mutually shared gaze. In *ICAT-EGVE*, pp. 197–204, 2017. 1, 2, 8

[26] Y. Li, F. Lu, W. S. Lages, and D. Bowman. Gaze direction visualization techniques for collaborative wide-area model-free augmented reality. In *Symposium on spatial user interaction*, pp. 1–11, 2019. 2

[27] S. Morrison-Smith and J. Ruiz. Challenges and barriers in virtual teams: a literature review. *SN Applied Sciences*, 2(6):1–33, 2020. 1

[28] N. Norouzi, A. Erickson, K. Kim, R. Schubert, J. LaViola, G. Bruder, and G. Welch. Effects of shared gaze parameters on visual target identification task performance in augmented reality. In *Symposium on Spatial User Interaction*, pp. 1–11, 2019. 2

[29] C. Schnier, K. Pitsch, A. Dierker, and T. Hermann. Collaboration in augmented reality: How to establish coordination and joint attention? In *ECSCW 2011: Proceedings of the 12th European Conference on Computer Supported Cooperative Work, 24-28 September 2011, Aarhus Denmark*, pp. 405–416. Springer, 2011. 2

[30] H. Shah, R. Shah, S. Shah, and P. Sharma. Dangerous object detection for visually impaired people using computer vision. In *2021*

International Conference on Artificial Intelligence and Machine Vision (AIMV), pp. 1–6. IEEE, 2021. 3

[31] J. Terken and J. Sturm. Multimodal support for social dynamics in co-located meetings. *Personal and Ubiquitous Computing*, 14(8):703–714, 2010. 2

[32] D. Ungureanu, F. Bogo, S. Galliani, P. Sama, X. Duan, C. Meekhof, J. Stühmer, T. J. Cashman, B. Tekin, J. L. Schönberger, et al. Hololens 2 research mode as a tool for computer vision research. *arXiv preprint arXiv:2008.11239*, 2020. 2, 5

[33] V. Weistroffer, A. Paljic, P. Fuchs, O. Hugues, J.-P. Chodacki, P. Ligot, and A. Morais. Assessing the acceptability of human-robot co-presence on assembly lines: A comparison between actual situations and their virtual reality counterparts. In *The 23rd IEEE International Symposium on Robot and Human Interactive Communication*, pp. 377–384. IEEE, 2014. 3

[34] E. Wilson, A. Ibragimov, M. J. Proulx, S. D. Tetali, K. Butler, and E. Jain. Privacy-preserving gaze data streaming in immersive interactive virtual reality: Robustness and user experience. *IEEE Transactions on Visualization and Computer Graphics*, 2024. 3

[35] Worksafe and partners. Safe & sustainable recycling: Protecting workers who protect the planet, 2015. Accessed: 2024-08-21. 3

[36] C. Xu, N. Han, and H. Li. A dangerous goods detection approach based on yolov3. In *Proceedings of the 2018 2nd International Conference on Computer Science and Artificial Intelligence*, pp. 600–603, 2018. 3

[37] Y. Zhang, K. Pfeuffer, M. K. Chong, J. Alexander, A. Bulling, and H. Gellersen. Look together: using gaze for assisting co-located collaborative search. *Personal and Ubiquitous Computing*, 21:173–186, 2017. 1, 2