INTERACTION AND VISUALIZATION DESIGN CONSIDERATIONS FOR GAZE-GUIDED COMMUNICATION IN COLLABORATIVE EXTENDED REALITY

Luca Brägger[1], Louis Baumgartner[1], Kathrin Koebel[1], Joe Scheidegger[2], Arzu Çöltekin[1*]

[1]Institute of Interactive Technologies, University of Applied Sciences and Arts Northwestern Switzerland, Switzerland
[2]SBB Center of Competence Extended Reality, Bern, Switzerland

Commission IV WG IV/9

KEY WORDS: Extended reality, interaction design, visualization design, gaze interaction, eye movements, navigation

ABSTRACT:

There is evidence in literature that collaborative work while using digital tools could benefit from visualizing the real time eye movements of a selected participant, or possibly, several participants. In this study, we examine alternative gaze interaction and visualization design prototypes in a digital collaboration scenario, in which assumed collaboration environment is a co-located mixed reality environment. Specifically, we implemented a virtual pointer as a baseline, and representations of gaze as a line, a cursor, and an 'automated line' where the line and cursor are automatically alternated based on occlusion detection. These prototypes are then evaluated in a series of usability studies with additional exploratory observations for a spatial communication scenario. In the scenario participants either describe routes to someone else or learn them from someone else for navigational planning. In this paper we describe the alternative interaction design prototypes, as well as various visualization designs for the gaze itself (continuous line and dashed line) and the point of regard (donut, dashed donut, sphere, rectangle) to guide collaboration and report our findings from several usability studies (n=6). We also interviewed our participants which allows us to make some qualitative observations on the potential function and usefulness of these visualization and interaction prototypes. Overall, the outcomes suggest that gaze visualization solutions in general are promising approaches to assist communication in collaborative XR, although, not surprisingly, how they are designed is important.

1. INTRODUCTION AND BACKGROUND

Arguably a key limitation in the current head-mounted extended reality (XR) systems is that they are optimized for a single user. While there are many efforts towards multi-user scenarios, the lack of effective tools that enable collaboration in MR is a severe limitation, because humans live, learn and work in social contexts, in contact with each other. On the other hand, in the context of collaborative work, the benefits of interaction in three-dimensional space are well understood (further elaborated in the next section), and therefore opportunities offered by XR technology are exciting. Compared to the now-traditional (even de-facto during Covid-19 outbreak for many professions) video communications, the idea of seamless embodied interaction in XR bears immense potential. It promises life-like interactions *e.g.,* being able to show what is right in front of you, passing an object —even virtually— to another person in a meeting, or having eye contact with our collaborators (even via avatars) are all exciting possibilities. However, while promising, such interactions in XR today are at best clumsy, and realistically, they are not yet here. In this paper we explore a specific interaction concept in the context of collaborative work: We examine how to visualize someone's gaze and design interactions with this gaze representation. We visualize gaze in three different ways (as a line, as a cursor and as an automated line) and compare them to a more traditional virtual 'pointer' metaphor as a baseline. Our contributions in this paper include a brief review of the interdisciplinary literature; concept development and prototype implementations of the abovementioned concepts that have been implemented based on user-centered design principles in multiple design iterations; and series of usability studies with six participants along with exploratory trials, including qualitative user feedback assessing our prototypes.

1.1. Collaboration

While we intuitively understand the term collaboration, it is important to remember that there are different forms of **collaboration**, such as between organizations, groups, or teams. Each offer advantages and disadvantages for individual team members and organizations (Landy and Conte, 2016). For individual team members, it has been demonstrated that collaboration can have an overall positive effect not only on extrinsic but also on intrinsic motivation (Carr and Walton, 2014), it can promote social ties (Baumeister and Leary, 1995), and help achieve common goals that would not be possible on their own (Asch, 1952; as cited in Rozin, 2001). Furthermore, the interest and motivation of people in complicated or sophisticated tasks can increase through collaboration (Isaac et al., 1999). For organizations, teamwork can save considerable time by, for example, parallelizing otherwise sequential activities where appropriate. Innovation and creativity can also be increased by combining a variety of ideas from team members, and necessary information can be obtained in teams quicker, enabling organizations to deliver products of higher quality more efficiently and effectively, and allowing knowledge exchange (Mohrman et al., 1995). Successful

---

[*] corresponding author arzu.coltekin@fhnw.ch

https://doi.org/10.5194/isprs-annals-V-4-2022-205-2022

collaboration requires careful team composition, training, motivation and the evaluation of team performance (Landy and Conte, 2016). While countless factors can influence team performance, most decisive appear to be trust, communication, leadership, and clear objectives (Bennett and Gadlin, 2012; Terveen, 1995). A question we are facing today is to what extent our digital tools can meet these requirements, and what must be done to build tools that facilitate trust, enable smooth communication between team members, enable leadership and convey the objectives of the work clearly.

1.2. Collaborative extended reality

As mentioned earlier, XR offers exciting potential to facilitate collaborative work via digital means. Since XR can mean many things, we will provide a brief definition of it first. Milgram and Kishino's (1994) well-known taxonomy for XR technologies presents a Reality-Virtuality Continuum (Milgram and Kishino, 1994) (Figure 1). The continuum places reality on the one end and the full virtuality on the other, placing mixed reality in between with some nuance. Today, this continuum is broadly termed XR, which contains augmented reality (AR), mixed reality (MR) and virtual reality (VR). Perhaps one term that changed in its definition over time is MR: Today –driven by industry, and picked up by many major actors in the XR scene– MR came to mean specifically the cases in which virtual objects are spatially registered to their physical locations in the world, often real time. On the other hand, AR can be a projection of the virtual object anywhere without a spatial reference. This definition deviates from Milgram and Kishino's (1994) classification (as shown in Figure 1), as it essentially puts MR as a sub-class of AR. According to this understanding, one can think of every MR as AR ('spatially registered AR'), but every AR will not be MR. An in-depth examination of the terminology on XR technologies, including a longer explanation of Milgram and Kishino (1994)'s classification, along with a current state of the art and research challenges can be found in recent publications (Çöltekin et al., 2020a, 2020b). Despite the confusion on the terms AR and MR, overall, definitions of XR and its sub-categories are relatively well understood. On the other hand, definition and requirements of a *collaborative* XR are still in their early stages.

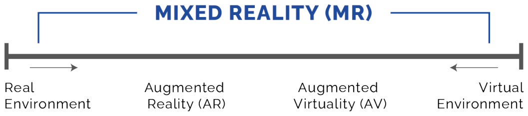


Figure 1. Milgram et al.'s (1994) seminal reality-virtuality continuum. Figure redrawn from Milgram (1994).

An implementation of collaborative XR requires understanding how collaboration works in a new spatiotemporal context. Since the early days of XR, a key argument to distinguish XR from other display systems has been that it can convey a *sense of presence* in a variety of ways (Sherman and Craig, 2003). In collaboration, sense of presence might be particularly important, as it can guide the attention of the participants to each other (Ens et al., 2019). A more natural depth perception provided in XR systems compared to 2D displays might lead to a stronger sense of presence. Perhaps partly because of these features, XR systems appear to facilitate more interactions (more cooperation, equal participation) among the participants than in video conferencing (Anton et al., 2018; Müller et al., 2017; Pan et al., 2018). If it is well-designed, XR allows a person to use their natural abilities to interact with each other as well as *with* and *in* the space, *e.g.,* through avatars, and this can amplify the sense of presence by creating the feeling that location-distributed participants are in the same room (Piumsomboon et al., 2017). In addition, XR can enable gestures and gaze-based features much more seamlessly and effectively than in other communication platforms, potentially facilitating implicit intention recognition along the way (Anton et al., 2018; Huang et al., 2015; Orts-Escolano et al., 2016; Pan et al., 2018; Tomasello et al., 2007). Even though gaze-based interactions and eye contact in XR are far from established, and their benefits are debated (Li et al., 2016), there is some early evidence that intention recognition might be already feasible (Akkil et al., 2016; Piumsomboon et al., 2017).

Mixed reality (MR) as a special type of XR, appears particularly promising for collaborative work with real world applications since virtual objects can be shown in their real world contexts (Billinghurst and Kato, 1999). There is considerable research on collaborative MR on various domains, e.g., in archeology (Benko et al., 2004), industrial product design (Ong and Shen, 2009), space operations (Fairchild et al., 2016), dancing (Zhou et al., 2019), learning (Giraudeau et al., 2019), 'time travel' (Pulver et al., 2020), etc. Many of these studies cite MR's benefits and how it improves collaborative processes and desired outcomes in the given context. While awareness cues and virtual objects as spatial cues have been examined in previous work, gaze-related work is rare, specifically, interaction and visualization design for using gaze as a collaborative cue in the context of navigation learning is largely unexplored. In navigation and other spatial contexts, eye movements have been explored in many different ways, e.g., for better understanding the cognitive processes involved in visuospatial information processing (Çöltekin et al., 2017), for examining the potential of gaze-contingent displays (Bektaş et al., 2019), or as a complimentary metric in usability studies (Çöltekin et al., 2009), etc. In the case of XR, along with other tracking technologies (Huesser et al., 2021), the use of gaze in any of these contexts is rapidly taking off, even though one can consider these efforts still in their early days.

1.3. Our study

Due to its importance in anticipating the intention of collaboration partners, we decided to examine gaze-guided communication in this study. Here we work a simulated gaze, i.e., gaze is estimated based on head movements and project from an assumed cyclopean eye, also known as *head-gaze.* Previous work has demonstrated that eye-gaze is superior to head-gaze a variety of ways (Blattgerste et al., 2018; Špakov et al., 2019). However, when eye-gaze is not available, head-gaze is a reliable-enough proxy to estimate rough gaze direction, and in some cases, it is comparable to eye-gaze (Špakov et al., 2019). For our purposes (i.e., to examine the interaction and visualization design of the gaze and the point of regard in collaborative settings) it is not critical if we work with a head gaze or actual (eye) gaze. Also considering not all headsets have integrated eye trackers, we decided to start our experiments with several variants of the head-gaze. We implemented various head-gaze protoypes: a *gaze-cursor*, a *gaze-line*, an *'automated' gaze-line*, and a *virtual laser pointer.* In addition, we implemented four visualization variants for the point of regard, and two for the gaze itself (elaborated in Section 2.3). We evaluated these prototypes in a series of mini experiments for a spatial communication task where participants collaboratively described or learned routes. We created these interaction and visualization alternatives to assist the collaboration process and collected subjective feedback on these alternatives.

It is important to note that in this exploratory study, the focus is on the iterative design of the interaction and visualization prototypes, i.e., we do not seek to confirm hypotheses (but build new hypotheses) at this stage. Despite its exploratory nature, when we designed the study, based on the observations in the user-centered design iterations, we expected that; 1) automated gaze-line may lead to best performance and might be preferred most by the participants, 2) gaze-line might be superior to gaze-cursor in drawing (*i.e.,* require fewer lateral body movements), 3) laser-pointer might lead to higher performance than head-gaze solution and receive higher usability ratings due to people's familiarity with the tool, the more familiar use of hands for drawing, and the fact that eyes can be used more freely to look at each other in a co-located MR setting.

2. EXPERIMENTS

After a concept development phase, we implemented the prototypes, and conducted the studies in a co-located synchronous MR setting with HoloLens (1gen). Our co-located synchronous MR implementations were for a scenario inspired by a selected use case at the Swiss Federal Railways (SBB), developed after interviewing experts at the SBB. In this scenario, two actors collaborate on coordinating navigation, specifically, route learning (elaborated in section 0) using a city model, i.e., a fictitious 3D map virtually displayed on a tabletop. In the scenario, two actors collaborate: *the communicator* is an active participant who sends a message, and *the recipient* is mostly passive, following instructions with the goal to learn the route. During the collaboration, the communicator uses the provided awareness cues (i.e., our implementations of three gaze visualization/interaction prototypes, and a virtual pointer) to help describe a route and highlight individual buildings or objects as landmarks in MR.

As mentioned earlier, this was an exploratory study, nonetheless, we designed several mini-experiments as the first step for later hypothesis testing. With that in mind, our independent variables were four interaction designs (gaze-line, gaze-cursor, automated gaze-line, and laser pointer), two visualization design variants for the gaze-line (continuous-line and dashed-line), and four visualization design variants for the gaze-cursor (donut, dashed-donut, sphere, dashed-rectangle). These variables are illustrated in Figures 5, 6 and 7. Among these variables, comparable ones were grouped and tested in different sessions, each session is intended as a mini-experiment or a pilot for a future controlled study (detailed in the section 2.3. Procedure). Throughout the prototype development, we continuously collected user feedback from small groups qualitatively, and measured various aspects of usability along with further qualitative feedback. Our *dependent variables* in the user studies were as follows: the number of inclinations of the upper body (count), number of errors in the route drawing on the map, *system usability scale* (SUS) (Brooke, 1996) scores for two of the comparable prototypes, a preference questionnaire (i.e., subjective ratings on the design variants) and, open-ended interview questions.

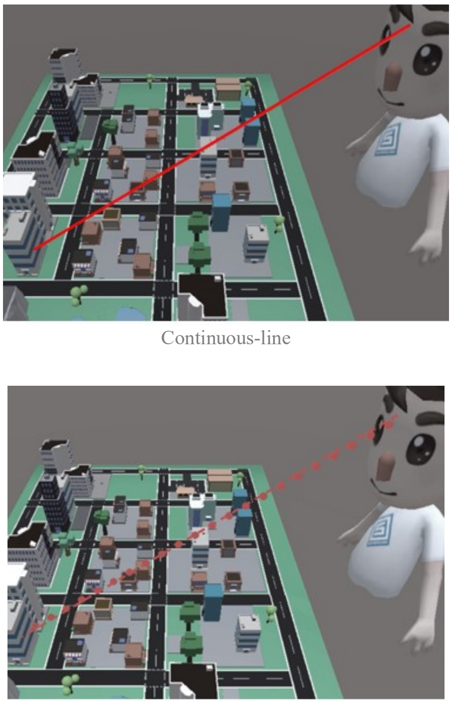
2.1 Participants

We recruited six participants, as it has been previously shown that six participants are sufficient in usability studies (Nielsen, 2012). Our scenario was aimed at expert users; therefore, our participants have a background in computer science, spatial sciences, or civil engineering. None of the participants had experience with MR displays before the study.

2.2. Prototype implementations (materials)

We implemented our MR prototypes for Microsoft HoloLens (1gen) (Zeller et al., 2021). The communicator (*i.e.,* collaboration partner who instructs the recipient) is simulated as an animated avatar in one of the sessions for control purposes, whereas in another session the participant takes the role of the communicator where the experimenter takes the role of the recipient. We adopted the avatar from the creator Supercyan (Supercyan, 2020), and animated it in Unity 3D LTS-Version 2018.4.10f1 (Unity QA, 2020). We also constructed a virtual city model, using the building elements 'Low Poly City' by the creator Viuletti (Viluetti, 2020). To ensure that the task is not 'too easy' for the visual prototypes, the first row of buildings in our city model intentionally consists of towers. In addition, for comparability, we set the buildings up symmetrically to each other in 13 sectors, *i.e.,* the routes were designed to be similar in their visual qualities and levels of difficulty. All prototypes were built with Unity 3D LTS-Version 2018.4.10f1 (Unity QA, 2020), and for the head-gaze prototypes and related interactive elements, we used the Mixed Reality Toolkit (MRTK) from Microsoft (Microsoft, 2021). We varied the prototypes in interaction or visualization design. Below we introduce each prototype in further detail.

Prototype 1: Gaze-line. In the gaze-line prototype, the communicator's gaze is visualized by two variants of red line (continuous vs. dashed), which moves with the communicator's head movement. The starting point of the line is between the communicator's eyes (*i.e.* the 'cyclopean eye'). The end point of the line (point of regard) is where the communicator's gaze hits the surface of a virtual object. In this prototype, the line is permanently displayed. If no virtual surface is being viewed, the line has a fixed length. Gaze-line has been implemented in two variants, continuous, and dashed-line versions (Figure 2). A continuous line representing the gaze gives precise information



Continuous-line

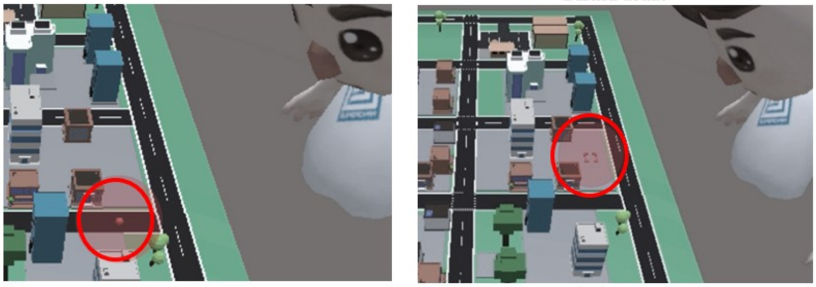
Dashed-line

Figure 2. An illustration of two gaze-line variations.

about the direction and the position of the point of regard serving as a good baseline. On the other hand, the head-gaze has its point of origin between the eyes, therefore it is unfavorable from the communicator's perspective if a whole line is displayed as it can occlude their view. For this reason, we experimented with a dashed-line, which is a simplified (visually less complex) visualization and might be less invasive for the task as it would occlude less of map. With similar reasoning, in the next prototypes, we worked on the ideas of showing only a cursor (gaze-cursor), and a combination of cursor and line (automated gaze-line).

Prototype 2: Gaze-cursor. In the gaze-cursor prototype, user's point of regard is displayed without a line in four variations (Figure 3). The rotation of the cursor is always aligned with the surface on which it rests. This is done by calculating a normal vector to the surface. The cursor is permanently displayed and moves with the movement of the communicator's head. To better understand the potential effects of different visualization types, we implemented four design variants for the gaze-cursor: *donut*, *dashed-donut, sphere, dashed-rectangle* (Figure 3). The visualization designs shown in Figure 3 were similarly motivated by occlusion / visual complexity arguments as explained above with the gaze-line prototypes. In the case of rectangular design, we wanted to explore if the shape of a rectangle would be rated differently than the circular/spherical shapes, because rectangle is similar to the objects in the background in an urban space, making it more coherent with the environment, thus might increase desirability. On the other hand, it might be less salient due to its similarity to the objects at the background. We displayed the gaze-cursor variants in red color (similarly to the gaze-line prototypes) because it provided a high contrast to the background colors therefore it did not have strong visibility issues due to matching colors.

Donut Dashed-donut



Sphere Dashed-rectangle

Figure 3. Visualization design variants for the gaze-cursor.

Prototype 3: Automated gaze-line. The automated gaze-line is a combination of the gaze-line and the gaze-cursor which algorithmically switches between these two modes. The communicator's point of regard is represented by a gaze-cursor (in this prototype, the donut variant), however, if the cursor is occluded by a virtual object, the gaze-line is automatically switched on. When the cursor becomes visible to the recipient again, the gaze-line disappears. The cursor is permanently displayed (i.e., only the line is adaptively shown) and both move with the head movement of the communicator.

Prototype 4: Virtual laser pointer. Here we use a 'laser pointer' metaphor to shoot a beam of light at a specific spot using a virtual device. The laser pointer is partially controlled by hand and has a switch-button called 'on/off' to switch between shooting a permanent beam when activated (nothing happens when deactivated). When activated, the communicator can redirect the beam to another point on the virtual object by using gaze and performing an air-tap. When holding the air-tap, the beam follows the gaze direction and locks at the last position when the air-tap is released. In addition, the laser pointer has a switch-button we called *follow me*. If *follow me* is activated, the laser pointer follows the user and remains in their field of vision. When deactivated, it remains in place. The position of the laser pointer (*i.e.,* the origin of the beam) can be moved directly at any time by initially gazing at it, then holding the air-tap and moving the hand. While moving the laser pointer, the endpoint of the beam remains on the same position (Figure 4).

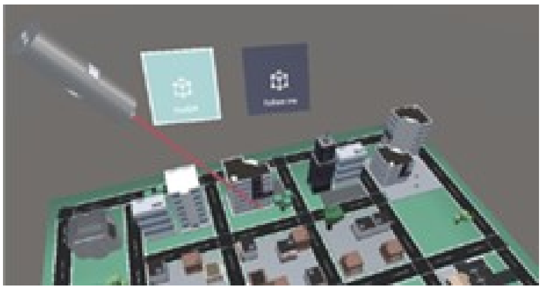


Figure 4. Visualization of the virtual laser pointer.

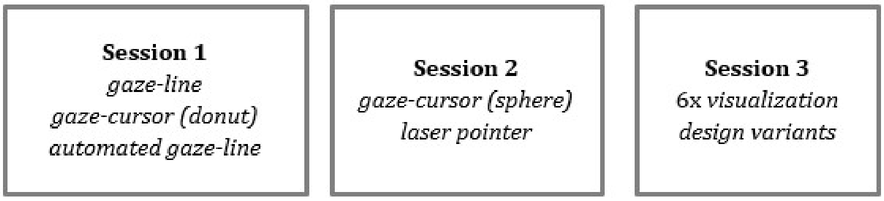


Figure 5. The overview of the independent variables in the three sessions of the experiment.

2.3. Procedure

Throughout the design process, we have collected qualitative feedback from small groups of users iteratively. Before the final experiments began, we briefed the participants, obtained informed consent, and offered training on how to operate the HoloLens. Throughout these experiments, broader context is that two people (the communicator and the recipient) exchange navigation information. They teach (instruct / describe) or learn a route from each other using a three-dimensional city model as a virtual table-top map, and various collaborative MR solutions described above to facilitate this interaction. The experiments were organized in three exploratory sessions (Figure 5). Below we further detail each session.

Session 1. The first session was designed to explore how well participants (all in the recipient role) can learn a route described by an avatar (the communicator). Participants were instructed to follow avatar's gaze and memorize the route, and two specific houses that were pointed by the communicator (with longer gaze durations). After each route, participants were asked to draw the memorized route using the three head-gaze variants (gaze-line, donut gaze-cursor, automated gaze-line) in a systematically rotated order. The communicator has shown a path similar in length and other qualities in each condition (the paths were not identical to counter against the learning effect). Participant was positioned on the opposite side of the city model from the avatar (Figure 6). During the experiment, participants remained in the same position to make sure that each participant's perspective was the same. We collected objective metrics (number upper body movements, number of errors in the route drawing) in the first session.

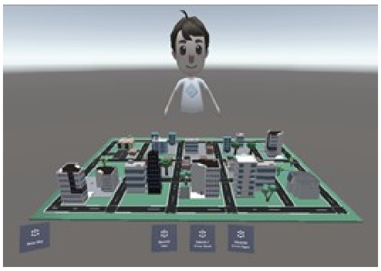


Figure 6. City model with avatar for Session 1.

Session 2. In this session, the primary goal was to explore user satisfaction with a head-gaze prototype vs. the laser-pointer. We selected sphere variant for the head-gaze (as a representative of the head-gaze implementations) because this variant is the most similar to the laser-pointer's point of regard visualization (i.e., the tip of the line). To have insights on the changing roles, participants took on the role of the communicator in this session whereas the moderator played the role of the recipient. Thus, they were asked to describe a route of their own choice, with start and end points were marked on a map shown in aerial perspective, which we speculated can be slightly more challenging, especially when communicating about landmarks (Lokka and Çöltekin, 2020). Participants were asked to choose a turn at each intersection to discourage a 'flight path' behavior, and they were asked to make eye contact with the recipient, every time they pass previously marked landmarks to encourage collaborative behavior, and to see how well they can get back to the path with different prototypes. In this session, we measured task duration, and SUS scores for the two prototypes.

Session 3. The goal of the third session was to get insights on the user preference among the design variants of the prototypes. We gathered subjective feedback (quantitative ratings and interview responses) from participants on the different design variants of gaze-line and gaze-cursor (Figure 3). We have shown participants these prototypes as videos, and asked them to subjectively rate the variants for their visual quality i.e., if participants liked the design visually; and for their imagined functionality i.e., if participants thought the variant would work well in the navigation related communication that they experienced in previous sessions.

3. RESULTS AND DISCUSSION

The responses and data from the Sessions 1 and 2, including the interview results, are summarized in Tables 1 and 2, and task duration per prototype is shown in Figure 7.

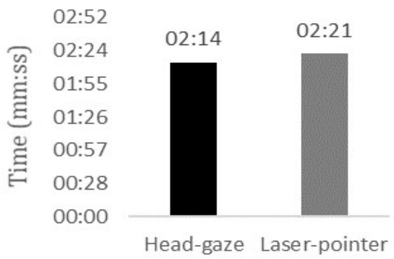


Figure 7. Average task duration in Session 2. Overall participants use slightly less time with the gaze cursor than with the virtual pointer (descriptive statistics only, n=6).

Figure 7 suggests that participants are slightly faster with the gaze cursor solution than our virtual pointer implementation, albeit with a very small difference. This slightly favorable result for the head gaze cursor in Figure 7 is consistent with the SUS score (Table 2), where we can see that gaze receives a considerably higher SUS score than virtual laser pointer (68.75 vs. 41.5). In the third session, among the four design options for the gaze-cursor, the donut-cursor received the best average rating (closely followed by the dashed-donut), both visually and functionally, whereas the continuous-line was clearly preferred to the dashed-line (Figure 8).

Table 1. Summary results from Session 1 where the participant is in the role of the recipient.

Measurement type	**Outcome variable**	**Summary of observations**
Performance metrics	Number of errors	On average, the *automated gaze-line* led to highest number of errors (12.7), followed by *gaze cursor* (9.6), whereas the *gaze-line* caused the fewest errors (7.1).
Observed behavior	Number of sideways body movements	The *gaze-cursor* has a significantly higher number of movements (8) than the *gaze-line* (4.7) and the *automated gaze-line* (3.8).
Subjective feedback	Most preferred	The *gaze-line* was most preferred by the participants (4 out of 6). Participants cited reasons as clarity and simplicity of following the gaze.
	Most criticized	The *gaze cursor* was criticized most often (5 out of 6). The most mentioned reason was the complete disappearance behind 3D objects. Switching the gaze-line on and off in the *automated gaze-line* was confusing for several participants.
	Suggestions for improvement and ideas	- gaze-line could be permanently displayed *with* the gaze-cursor - objects being looked at could be visually highlighted - important points could be marked, *e.g.,* with a pin or an outline

Table 2. Summary results from Session 2 where the participant is in the role of the communicator.

Measurement type	**Outcome variable**	**Summary of observations**
SUS	Test results	The result for the *gaze cursor* with an average score of 68.75 is much better than the *laser pointer* with an average score of 41.5.
Subjective feedback	Most preferred	The *laser pointer* was preferred more often than the *gaze cursor* in the context of collaboration (4 out of 6).
	Most criticized	The *gaze cursor* was often criticized for being more complicated to make eye contact with other persons. With the *laser pointer*, the frequent clicking to mark a spot can be tiring.
	Suggestions	*Gaze cursor* should provide the option to set additional markers.

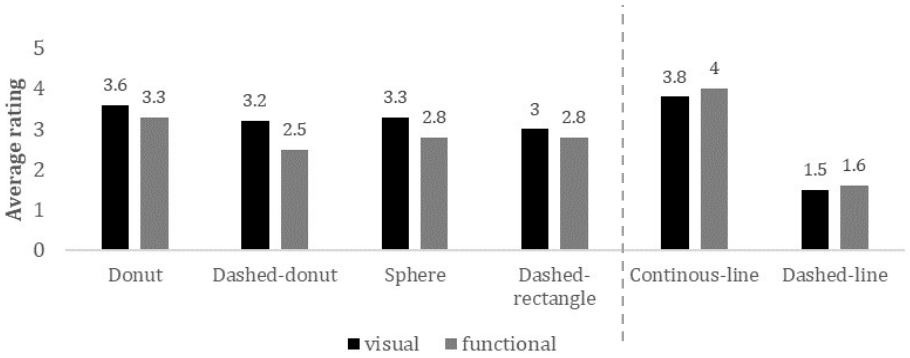


Figure 8. Participants' average ratings (1 "very bad" to 5 "very good") for the six visualization design options. First four are the point-of-regard visualizations of gaze, whereas the last two are visualizations of the gaze line. Participants rate the continuous donut-cursor best among the point of regard visualization options, and they rate continuous-line clearly higher than the dashed line for the gaze line visualization options.

Thus, qualitative results presented in Tables 1 and 2 both highlight that the gaze based solutions have potential, however, we can see that the interaction design matters. Figure 8 reflects the qualitative comments on Table 1 for the gaze visualization options, where participants rate the continuous line clearly higher than dashed line, and in Table 1, they stated a clear preference for the gaze-line, and committed fewest errors with it. We also see that among the point-of-regard visualization options, donut (which is also a continuous line) is rated highest for both aesthetics and (perceived) function, though the differences between donut and others are fairly small. Below we interpret these results in the light of the qualitative interviews and insights based on our observations.

Considering occlusion and visual complexity arguments, we expected that our participants would perform best with the *automated gaze-line* in the first session, where they were in the role of the recipient. Contrary to this expectation, participants committed the least errors with (and preferred) the *gaze-line* instead when they were asked to draw the route that they were asked to learn. In the interviews, participants stated that the gaze-line was the easiest for the task because it is more consistent and clearer compared to the other prototypes. On the other hand, some participants mentioned that the gaze-line was irritating, especially when they wanted to find the exact position on an object. We believe a cursor shown at the precise point of regard might solve this problem. This line of thinking is also supported by various positive statements about the automated gaze-line in the interviews where participants mention that combination of cursor and gaze-line offers better focus. However, the gaze-cursor alone was not convincing in the first session, participants committed more errors in their route drawing than with the gaze-line, and because it was partially occluded by the virtual objects, participants had to lean to the sides most often (Table 1).

The automated gaze-line, on the other hand, led to the worst results in route drawing, and the best results in the number of lateral body movements, suggesting that even though the main idea may be promising, its design and implementation needs further consideration. Participants liked the solution when the gaze-line is switched on (thus they had gaze line + gaze cursor at the same time). Automatically switching the gaze-line on and off creates unexpected motion in the peripheral vision, which is known to distract people, which causes them to lose track on the route and leads to confusion. In a future design, it is conceivable that gaze-line does not disappear completely but only increases or decreases in intensity (as expressed by transparency, or line thickness), and most importantly, motion during these changes need to be slow, smooth, and barely noticeable. Experimenting with different levels of transparency may also address the occlusion problem.

Among the design variants of the gaze-cursor, our participants rated the *donut* cursor best, and coincides with Microsoft's recommendation (Turner et al., 2022). However, ratings for the gaze-line are even higher (Figure 9). In sum, both the gaze-line and the gaze-cursor appears to offer an added value, thus we surmise that future representations should have elements of both. In our results the dashed-versions of either the cursor or the line were not popular. We believe this may be partially due to the background images that were visually rather complex (many colors, 3D viewed from an oblique view with some occlusion, containing many buildings and other details) and dashed versions were too subtle. In a more ‘rural’ environment (or any scene with less visual complexity), it is possible that the dashed versions work well and may be rated higher. A future study where the background is varied may allow for a more robust evaluation. Furthermore, if the cursor is ‘flexible’, i.e., if it can adapt to the shape of an object, that may better support perception in three-dimensional spaces. A gaze-based highlighting of the objects at the point of regard may also improve the visuospatial orientation of the viewer.

In Session 2, participants took the role of the communicator, and were tasked with describing a route by using a donut gaze-cursor *vs.* a laser pointer. We expected that participants would give the laser pointer a higher usability score (in the SUS) than the head-gaze prototype, and would evaluate it more positively in the interviews too. Surprisingly, the head-gaze prototype received a remarkably higher SUS score (68.75) than the laser-pointer (41.5), even though the participants’ statements in the interviews suggest that the laser pointer might receive greater acceptance for collaborative work. This too, might be a design problem: Participants unanimously expressed that it is more pleasant to make eye contact with the meeting partners, if the marker is detached from the head and remains in place. Controllers in virtual reality are the current de-facto standard, with better hand tracking, a laser-pointer metaphor supported by gaze visualizations may offer viable solutions. Allowing more training time with the laser-pointer than what we offered our participants (since air-tap is still a new way of interacting with information) might also be a good approach.

4. CONCLUSIONS AND OUTLOOK

Based on prototype implementations, we examined the use of MR in collaborative work in a co-located setting. Our interview results and usability studies give indications that a) gaze-based solutions are promising approaches, b) the elements of design (the ‘nitty gritty details’) do matter, *e.g.,* it indeed makes a difference if we use a continuous line or a dashed line when we represent gaze. Remaining quantitative results are exploratory, and as the exploratory studies are meant to facilitate, they gave us insights on the prototype implementations to design future controlled laboratory experiments. A clear next step is to implement a proper eye tracking alternative to the head-gaze solution based on the lessons learned in this preliminary set of experiments; and measure more on the collaboration related metrics (our focus here has been on usability) as well as factors such as learnability of the novel solutions and user fatigue. As we carry out the planned research, we envision that our findings will be relevant and useful for other domains such as education, behavior change, stress identification and response in emergency, as well as collaborative spatial decision making under time pressure or threatening situations.

ACKNOWLEDGEMENTS

We are grateful to our participants and reviewers. This project was partially supported by the SBB Center of Competence for Extended Reality and Innosuisse project 50321.1.

REFERENCES

Akkil, D., James, J.M., Isokoski, P., Kangas, J., 2016. GazeTorch: Enabling Gaze Awareness in Collaborative Physical Tasks, in: Proc. CHI EA ’16. ACM Press, SJ, CA, USA, pp. 1151–1158.

Anton, D., Kurillo, G., Bajcsy, R., 2018. User experience and interaction performance in 2D/3D telecollaboration. Future Generation Computer Systems 82, 77–88.

Asch, S.E., 1952. Group forces in the modification and distortion of judgments.

Baumeister, R.F., Leary, M.R., 1995. The need to belong: desire for interpersonal attachments as a fundamental human motivation. Psychological bulletin 117, 497.

Bektaş, K., Çöltekin, A., et al., 2019. GeoGCD: improved visual search via gaze-contingent display, in: Proc. of the 11th ACM Symposium ETRA ’19, ACM, Denver Colorado, pp. 1–10.

Benko, H., Ishak, E.W., Feiner, S., 2004. Collaborative mixed reality visualization of an archaeological excavation, in: Third IEEE and ACM International Symposium on Mixed and Augmented Reality. IEEE, pp. 132–140.

Bennett, L.M., Gadlin, H., 2012. Collaboration and team science: from theory to practice. Journal of Investigative Medicine 60, 768–775.

Billinghurst, M., Kato, H., 1999. Collaborative mixed reality, in: Proc. of the First International Symposium on Mixed Reality. pp. 261–284.

Blattgerste, J., Renner, P., Pfeiffer, T., 2018. Advantages of eye-gaze over head-gaze-based selection in virtual and augmented reality under varying field of views, in: Proc. of the Workshop on Communication by Gaze Interaction. pp. 1–9.

Brooke, J., 1996. SUS-A quick and dirty usability scale. Usability evaluation in industry 189, 4–7.

Carr, P.B., Walton, G.M., 2014. Cues of working together fuel intrinsic motivation. Journal of Experimental Social Psychology 53, 169–184.

Çöltekin, A., Brychtová, A., et al., 2017. Perceptual complexity of soil-landscape maps: a user evaluation of color organization in legend designs using eye tracking. International Journal of Digital Earth 10, 560–581.

Çöltekin, A., Griffin, A.L., Slingsby, A., Robinson, A.C., et al., 2020a. Geospatial information visualization and extended reality displays, in: Manual of Digital Earth. Springer, Singapore, pp. 229–277.

Çöltekin, A., Heil, B., Garlandini, S., Fabrikant, S.I., 2009. Evaluating the Effectiveness of Interactive Map Interface Designs: A Case Study Integrating Usability Metrics with Eye-Movement Analysis. Cartography and Geographic Information Science 36, 5–17.

Çöltekin, A., Lochhead, I., Madden, M., Christophe, S., et al., 2020b. Extended reality in spatial sciences: a review of research challenges and future directions. ISPRS International Journal of Geo-Information 9, 439.

Ens, B., Lanir, J., Tang, A., Bateman, S., Lee, G., et al., 2019. Revisiting collaboration through mixed reality: The evolution of groupware. International Journal of Human-Computer Studies 131, 81–98.

Fairchild, A.J., Campion, S.P., García, A.S., Wolff, R., et al., 2016. A mixed reality telepresence system for collaborative space operation. IEEE Trans. on Circuits and Systems for Video Tech. 27, 814–827.

Giraudeau, P., Olry, A., Roo, J.S., Fleck, S. et al., 2019. CARDS: a mixed-reality system for collaborative learning at school, in: Proc. of the 2019 ACM Int. Conf. on Interactive Surfaces and Spaces. pp. 55–64.

Huang, C.-M., Andrist, S., Sauppé, A., Mutlu, B., 2015. Using gaze patterns to predict task intent in collaboration. Front. Psychol. 6.

Huesser, C., Schubiger, S., Çöltekin, A., 2021. Gesture Interaction in Virtual Reality: A Low-Cost Machine Learning System and a Qualitative Assessment of Effectiveness of Selected Gestures vs. Gaze and Controller Interaction, in: Proc. of the INTERACT 2021, LNCS. Springer, Cham, pp. 151–160.

Isaac, J.D., Sansone, C., Smith, J.L., 1999. Other people as a source of interest in an activity. Journal of Experimental Social Psychology 35, 239–265.

Landy, F.J., Conte, J.M., 2016. Work in the 21st century: An introduction to industrial and organizational psychology. John Wiley & Sons.

Li, J., Manavalan, M.E., D'Angelo, S., Gergle, D., 2016. Designing Shared Gaze Awareness for Remote Collaboration, in: Proc. of the 19th ACM Conference on Computer Supported Cooperative Work and Social Computing Companion - CSCW '16. ACM Press, SF, CA, USA, pp. 325–328.

Lokka, I.E., Çöltekin, A., 2020. Perspective switch and spatial knowledge acquisition: effects of age, mental rotation ability and visuospatial memory capacity on route learning in virtual environments with different levels of realism. Cartography and Geographic Information Science 47, 14–27.

Microsoft 2021. Microsoft MixedRealityToolkit-Unity. https://bit.ly/36iC4Eh (accessed 23.03.2022)

Milgram, P., Kishino, F., 1994. A taxonomy of mixed reality visual displays. IEICE TRANSACTIONS on Information and Systems 77, 1321–1329.

Mohrman, S.A., Cohen, S.G., Morhman Jr, A.M., 1995. Designing team-based organizations: New forms for knowledge work. Jossey-Bass.

Müller, J., Rädle, R., Reiterer, H., 2017. Remote collaboration with mixed reality displays: How shared virtual landmarks facilitate spatial referencing, in: Proc. of the 2017 CHI Conference on Human Factors in Computing Systems. pp. 6481–6486.

Nielsen, J., 2012. How many test users in a usability study. Nielsen Norman Group 4.

Ong, S., Shen, Y., 2009. A mixed reality environment for collaborative product design and development. CIRP annals 58, 139–142.

Orts-Escolano, S., Rhemann, C., Fanello, S., et al., 2016. Holoportation: Virtual 3D Teleportation in Real-time, in: Proc. of the 29th UIST '16, ACM, Tokyo Japan, pp. 741–754.

Pan, Y., Sinclair, D., Mitchell, K., 2018. Empowerment and embodiment for collaborative mixed reality systems: Empowerment and Embodiment. Comput Anim Virtual Worlds 29, e1838.

Piumsomboon, T., Day, A., Ens, B., et al., 2017. Exploring enhancements for remote mixed reality collaboration, in: SIGGRAPH Asia 2017 Mobile Graphics & Interactive Applications. pp. 1–5.

Pulver, Y., Merz, C., Koebel, K., Scheidegger, J., Cöltekin, A., 2020. Telling engaging interactive stories with extended reality (XR): Back to 1930s in Zurich's main train station. ISPRS Annals of Photogrammetry, Remote Sensing & Spatial Information Sciences 5.

Rozin, P., 2001. Social psychology and science: Some lessons from Solomon Asch. Personality and Social Psychology Review 5, 2–14.

Sherman, W.R., Craig, A.B., 2003. Understanding virtual reality—Interface, application, and design. MIT Press One Rogers Street, Cambridge, MA 02142-1209.

Špakov, O., Istance, H., Räihä, K.-J., Viitanen, T., Siirtola, H., 2019. Eye gaze and head gaze in collaborative games, in: Proc. of the 11th ACM Symposium on Eye Tracking Research & Applications. pp. 1–9.

Supercyan, 2020. Character Pack: Free Sample - Asset Store https://bit.ly/3HMin4T (accessed 23.03.2022).

Terveen, L.G., 1995. Overview of human-computer collaboration. Knowledge-Based Systems 8, 67–81.

Tomasello, M., Hare, B., Lehmann, H., Call, J., 2007. Reliance on head versus eyes in the gaze following of great apes and human infants: the cooperative eye hypothesis. Journal of Human Evolution 52, 314–320.

Turner ('thetuvix') et al., 2022. Cursors - Mixed Reality https://bit.ly/3q9LadH (accessed 23.03.2022).

Unity QA 2020. Unity QA - LTS Releases [WWW Document], 2020. Unity. URL https://bit.ly/3zHCIWb (accessed 23.03.2022).

Viluetti, 2020. Low poly city from Viuletti - Asset Store https://bit.ly/33eNOpp (accessed 23.03.2022).

Zeller (' mattzmsft') et al., 2021. HoloLens (1st gen) hardware [WWW Document]. URL https://bit.ly/3zGjxfc (accessed 23.03.2022).

Zhou, Z., Márquez Segura, E., Duval, J., John, M., Isbister, K., 2019. Astaire: A Collaborative Mixed Reality Dance Game for Collocated Players, in: Proc. of the CHI PLAY '19: ACM, Barcelona Spain, pp. 5–18.