Measuring and Comparing Collaborative Visualization Behaviors in Desktop and Augmented Reality Environments

Michael Kintscher
Arizona State University
Tempe, Arizona, USA
mkintsch@asu.edu

Jinbin Huang
Arizona State University
Tempe, Arizona, USA
jhuan196@asu.edu

Anjana Arunkumar
Arizona State University
Tempe, Arizona, USA
aarunku5@asu.edu

Ashish Amresh
Northern Arizona University
Flagstaff, Arizona, USA
ashish.amresh@nau.edu

Chris Bryan
Arizona State University
Tempe, Arizona, USA
cbryan16@asu.edu

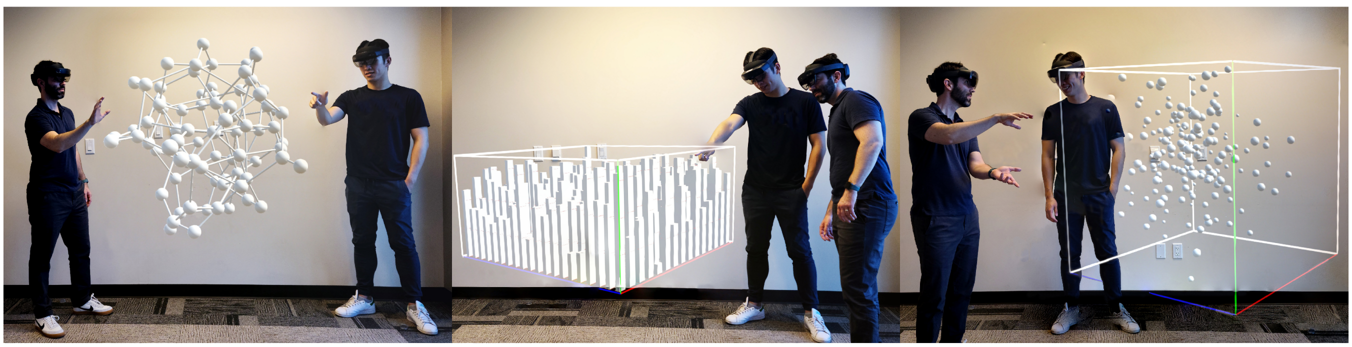

Figure 1: Examples of collaborative positional arrangements encoded during the study across the three visualization techniques used: (left) Same Space on a node-link diagram, (middle) Separate Space on a bar chart, and (right) Mixed Space on a scatter plot.

ABSTRACT

Augmented reality (AR) provides a significant opportunity to improve collaboration between co-located team members jointly analyzing data visualizations, but existing rigorous studies are lacking. We present a novel method for qualitatively encoding the positions of co-located users collaborating with head-mounted displays (HMDs) to assist in reliably analyzing collaboration styles and behaviors. We then perform a user study on the collaborative behaviors of multiple, co-located synchronously collaborating users in AR to demonstrate this method in practice and contribute to the shortfall of such studies in the existing literature. Pairs of users performed analysis tasks on several data visualizations using both AR and traditional desktop displays. To provide a robust evaluation, we collected several types of data, including software logging of participant positioning, qualitative analysis of video recordings of participant sessions, and pre- and post-study questionnaires including the NASA TLX survey. Our results suggest that the independent viewports of AR headsets reduce the need to verbally communicate about navigating around the visualization and encourage face-to-face and non-verbal communication. Our novel positional encoding method also revealed the overlap of task and communication spaces vary based on the needs of the collaborators.

CCS CONCEPTS

• **Human-centered computing** → *Human computer interaction (HCI)*; **Mixed / augmented reality**; **Collaborative interaction**; **User studies**;

KEYWORDS

Augmented reality, Visualization, Co-located collaboration

ACM Reference Format:
Michael Kintscher, Jinbin Huang, Anjana Arunkumar, Ashish Amresh, and Chris Bryan. 2023. Measuring and Comparing Collaborative Visualization Behaviors in Desktop and Augmented Reality Environments. In *29th ACM Symposium on Virtual Reality Software and Technology (VRST 2023), October 09–11, 2023, Christchurch, New Zealand*. ACM, New York, NY, USA, 11 pages. https://doi.org/10.1145/3611659.3615691

VRST 2023, October 09–11, 2023, Christchurch, New Zealand

ACM ISBN 979-8-4007-0328-7/23/10…$15.00
https://doi.org/10.1145/3611659.3615691

1 INTRODUCTION

Visualization is widely employed for analyzing, reasoning, and making decisions about data, but as data-driven problems become larger and more complex, support and understanding for collaborative visualization is becoming increasingly important [14, 22, 37].

Immersive modalities such as augmented reality (AR) are increasingly being used for 3D visualization to take advantage of semantics like stereoscopic rendering. Unfortunately, there is relatively little research with visualization-based user studies in AR, particularly ones that consider collaborative visual analysis. For example, collaborative AR is one of the least-studied aspects of the field, consisting of only 1.7% of such papers published in ISMAR across the 2010s [43]. Such studies are important for establishing empirical guidelines and best practices, meaning there are significant research questions in collaborative AR visualization contexts, such as how to support team-based analysis, reasoning, hand-off, and decision-making. Additionally, consistent, quantitative evaluation methods are needed to increase the rigor of such studies.

In particular, we are interested in following research question (RQ): "***How do dyads (two-person teams) collaborate when performing visual analysis of 3D datasets in AR?***" We investigate this question in two primary ways: (§3) We first propose a novel positional coding method to quantify the collaborative coupling of co-located dyads for an AR contexts based on measuring the overlap of their *task space* and *communication space*. This method, based on prior work by Tang et al. [49], and updated for head-mounted display (HMD) devices, can be used to promote a standardized analysis of collaboration behaviors. (§4–5) We then design, conduct, and analyze an experiment where dyads perform both closed and open-ended analysis tasks on visualizations of 3D datasets, in both AR and with a desktop-based computer (the latter modality serving as a comparative baseline). We conduct an extensive coding of participant actions and communications to analyze collaboration between dyad team members in the AR and desktop modalities.

Our study results provide nuanced insights into how participant behavior changes in AR vs. desktop scenarios — e.g., participants gesture and observe each other significantly more when collaborating in AR, suggesting that non-verbal movement and view-sharing between participants play a fundamentally different role in achieving a shared understanding and sensemaking among collaborators in AR (compared to traditional desktop displays).

We also demonstrate our novel positional coding method on the AR trials from the experiment. In particular, our method shows that participants frequently synchronize their views for communication, challenging the idea that overlapping communication and task spaces are inherently beneficial to users. Instead, the overlap (or lack thereof) seems to depend on the collaborative behaviors participants engage in, and those behaviors change frequently over the course of collaborating on a single visual analysis task.

Based on analyzing our experiment's results (including the application of the positional coding method), we propose several recommendations for dyad-based analysis tasks that use 3D visualization, and identify future research questions in the domain.

2 RELATED WORK

2.1 Communication.

Computers create an artificial separation between the "task space" and "communication space" [21], illustrating the importance of the overlap of these spaces in evaluating communication between collaborating users in computer supported cooperative work (CSCW) systems. Billinghurst and Kato define a "task space" as "*the shared workspace*" where tasks are performed, a "communication space" as "*the common interpersonal space*" where collaborators communicate with one another, and state the former is often a subset of the latter in face-to-face conversation. Applying the work of Ishii et al. [23], they say CSCW systems introduce seams when the task space is not a subset of the communication space, as users are forced to switch between the two [4]. Collaborative visualization systems should be designed to remove such seams [23], so modalities that encourage an overlap of task and communication spaces are desirable. Our study considers this overlap in both desktop and AR modalities, shown in Figure 2, and the novel method we propose for qualitatively encoding user positions in collaborative AR provides an empirical means to quantitatively measure this overlap.

"Conversational grounding" is the development of mutual understanding between conversational participants. Visual information is a vital part of collaborative communication because it helps participants gain situational awareness (an understanding of the state of the space) and establish conversational grounding [27]. Gergle et al. observed participants relying more on verbal communication as their shared view of the task space decreased, and concluded that showing participants what the other is doing is not enough; both participants need a shared understanding of what the other can see [19]. When collaborators could see a shared workspace but not each other, Ou et al. observed a strong connection between the difficulty of communicating about the task and a need to rely on vocal communication to establish conversational grounding [39].

2.2 Collaborative Augmented Reality.

Collaborative software (i.e., groupware) is often organized using a a space-time matrix [15, 24] based on whether collaborators are spatially *co-located* or *remote*, and if interaction occurs *synchronously* or *asynchronously*. Schmalstieg et al.'s work during the 1990s [42, 46] identified five key advantages to collaborative mixed reality: Virtuality, Augmentation, Cooperation, Independence, and Individuality [4]. More broadly, several studies have found AR contributes to collaboration [2, 4, 7, 11, 48], and has advantages for co-located, synchronous collaboration [31, 32, 34, 47]. However, a recent survey of AR research by Dey et al. [10] found that most collaborative user studies examined remote collaboration; co-located collaboration was identified as an opportunity for future research.

Most AR systems are visualization-based, designed to allow users to visualize, annotate, and inspect 3D models collaboratively [38]. Collaborating AR users are better able to coordinate actions when sharing a common point of view because they can rely on physical reference points in their shared environment [9, 36]. Additionally, placing the workspace between users reduces the amount of verbal communication to accomplish collaborative tasks, as it places the task space as a subset of the communication space like in natural, unmediated conversation [26]. The positional encoding technique we propose provides a consistent way to measure this placement.

2.3 Collaborative and 3D Visualization.

Collaborative visualization occupies a unique design and research space where success requires combining aspects of data analysis, teaming, groupware, and perception [22]. Prior studies have examined various workflows for collaborative teaming (e.g., [33, 51, 55]), but few have investigated communication and collaboration behaviors (particularly for dyads) for visualizations in AR.

3D visualizations are commonly used in both individual and collaborative work in stereoscopic environments (e.g., AR, VR, and CAVEs) [37]. While 3D visualization has well-known drawbacks (such as occlusion and perspective distortion) [37], it's been shown to better support certain tasks and sensemaking mechanisms: 3D visualizations improve a sense of context when analyzing data [45], certain types of tasks (e.g. tasks requiring orientation, navigation, or viewing a larger context) are better suited for 3D visualizations [50], and the higher levels of immersion afforded by stereoscopic 3D visualization aids users in evaluating presented data [41].

2.4 Collaboration Styles

Tang et al. [49] conducted a pair of user studies with participants engaging in exploratory information-finding on tabletop displays to observe collaborative coupling and proposed four collaboration styles ranging between highly engaged and highly disengaged. These styles were defined in part by encoding communication patterns and the positioning of users at fixed locations around a table relative to their partners. Isenberg et al. [22] later extended this work and proposed additional collaboration styles. These styles were focused on tabletop displays, and to the best of our knowledge, similar studies investigating collaboration styles using AR for data visualization and analysis have not yet been conducted. To address this, our study uses the similar measures to Tang et al. and builds upon their work to propose an updated method (see Section 3) to similarly classify observed behaviors into collaboration styles within the context of HMDs.

3 MEASURING SPACE OVERLAP

One significant motivation for this work is the current lack of a quantitative method for measuring the communication and task space overlap between co-located, synchronous collaborators when such collaborators are able to freely move about a workspace.

The position coding used by Tang et al. [49] serves as the inspiration for the novel method of position coding used in this study. Tang et al. classify the position of two participants each in one of six spaces around a tabletop visualization and encode the collective arrangement of the participants based on their relative positions to each other. This technique was effective, but was limited to participants each occupying one of several predefined locations; free-form participant movements across a space cannot be encoded. Our positional coding method overcomes this limitation while similarly using the relative positions of participants, and goes further to provide a means to measure the relative degree of overlap of the communication and task spaces noted by Billinghurst et al. [5].

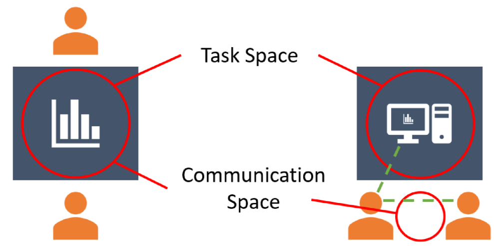

Figure 2: Shown is the overlap of communication and task spaces in two positional arrangements of co-located, synchronous collaborators. The left has overlapping spaces. The right has separate (non-overlapping) spaces and corresponds with common desktop setups, such as the setup in this study.

3.1 Locating Task and Communication Space

To adapt the positional arrangement for a free-form AR space in a way that measures overlap of the communication and task spaces, consider the 2D area formed by projecting a top-down view of the work space into a plane. (Figure 3 illustrates the example within the 15×15 foot space used in our study.) The visualization and each participant are represented by points on this plane. A line can be drawn through each point to form a triangle.

The task space exists centered around the visualization, and is defined by the space taken up by the visualization and any interfaces for interacting with it. Although the exact shape of the communication space is unknown, a few reasonable assumptions can be made about its location in space. The communication space is used for sharing communication cues such as gaze, gestures, and non-verbal behaviors [5]. Two observations can be made: (1) Each of these cues depend on the physical reach of the person communicating, thus such gestures can only be made in the area within reach of the communicator's body, and (2) these cues must be observed to be part of the communication. Considering these two observations, the communication space must include the area around each participants' body as well as the line-of-sight between them. While the exact size and shape of the communication space will vary with differing body shapes, sizes, and motor ability, the space itself will be largely confined to an ellipsoid defined by the position of the two people engaging in the communication. A line drawn between the two participants will thus approximate the center of this space.

3.2 Defining and Measuring Space Overlap

The degree of overlap of the two spaces from the perspective of each participant can thus be defined as proportional to the distance between the line between the participants and the center of the chart. Furthermore, the angle formed by the line between participant A and participant B and the line between the same participant A and the visualization is proportional to that distance and thus can be used to examine the overlap of the two spaces for participant A (and vice-versa for participant B). It is important to note that the two angles, one from each participant's point of view, are completely independent despite belonging to the same triangle, as no other restrictions on the third angle nor the length of any sides are placed. Thus one participant's degree of overlap does not impact the other participant's degree of overlap (given unlimited space).

The relative amount of overlap can be measured by comparing the angles between participants over time. This will provide a sense of the movement of the participants' spaces and a continuous measure of the relative space overlap as the participants collaborate. Changes in this measure can be used as evidence of a shift in collaborative behavior, similarly to Tang et al.'s analysis [49]

3.3 Positional Encoding

To aid using this measure as a data point in classifying behaviors, the perspective of a participant can optionally be categorized into *overlapping* or *non-overlapping* spaces, analogous to the two setups depicted in Figure 2 and described by Billinghurst et al. [5]. Independently applying these two categories to each participant leads to three possible arrangements (unique combinations): (1) *Same Space* - where both participants have overlapping spaces, (2) *Mixed Space* - where only one participant has overlapping spaces and the other has non-overlapping spaces, and (3) *Separate Space* - where both participants have non-overlapping spaces.

Each participant's angles are encoded into these categories by selecting a threshold angle. This threshold depends on the size of the task space and the nature of the tasks being performed, but the technical limitations of the hardware can be a practical upper-bound. HMD's have a limited field-of-view, so the horizontal viewing angle (angle formed between the two vertical edges of the device's display and the center point between the user's eyes) can serve as such an upper-bound; if the angle between the participant's partner and the visualization were any greater, the HMD would be incapable of rendering the visualization for the participant. Thus, any angles greater must be non-overlapping views. We chose 43.3 degrees to correspond with HoloLens 2's threshold (used in our study). This threshold was selected because it represents the horizontal viewing angle of the HoloLens 2 (calculated from the reported 52 degree diagonal viewing angle on a 3:2 area [35]), and thus is the greatest possible angle at which a participant could simultaneously see the visualization and their teammate without turning their head.

While we only consider AR HMDs in this paper (specifically HoloLens 2), this same technique can be extended for all types of AR devices with appropriate modification. Determining the upper-bound threshold for handheld devices (HHDs) is more complicated, as the position of the device's display is not fixed relative to the user's eyes; the user may hold the device close to their face, or at arm's length. Projector displays do not have an upper-bound threshold, but other thresholds dependent on the participant's natural (non-augmented) field-of-view will still apply. Once a threshold is derived, the remainder of the technique works the same for all devices. A threshold value should be chosen that makes sense for the tasks and visualization being used, with the upper-hound serving as a fallback if a more-precise threshold is not known.

4 STUDY DESIGN

To investigate the research question defined in §1, we propose three related hypotheses, and design and run a within-subject experiment to test them, where dyads are shown visualizations in AR (and on a desktop baseline) and must perform common visual analysis tasks. Our hypotheses center around three types of actions commonly used in collaborative sensemaking and analysis [22, 49]: (1) physical gestures, (2) verbal communication, and (3) spatial proximity.

4.0.1 Physical Gestures. Kiyokawa et al. [26] found that teammates communicated more with gestures than verbally when a collaborative visualization was placed between them, as compared to the side of both. When an AR visualization is placed in an arena, users can freely move around it. In contrast, teammates sitting in front of a desktop computer sit side by side. We thus hypothesize:

H1: *Participants will use more gestures when communicating while using AR than while using the desktop.*

H2: *Participants will use more gestures while using AR while in the Same Space arrangement than in the other arrangements.*

4.0.2 Verbal Communication. Gergle et al. [19] observed collaborators relying more on verbal communication as the amount of the workspace they visually shared decreased. When collaborators stand apart from one another, they have different points of view of the workspace; the amount they visually share decreases. Compared to AR, the desktop forces visually sharing a significant amount of the workspace. We thus further hypothesize that:

H3: *Participants will verbally communicate more when using AR than when using the desktop.*

H4: *Participants will verbally communicate more in AR while in the Same Space arrangement than in the other arrangements.*

4.0.3 Spatial Proximity. Billinghurst and Kato [4] suggests that overlapping the communication and task space is beneficial to users. Fussell et al. [18] found that collaborators focus their gaze equally as much on their partner (communication space) as they do on their tools and task (task space). We thus hypothesize that:

H5: *Participants will spend more time in the Same Space arrangement than in either of the other two arrangements.*

4.1 Experimental Design

The experimental design was within-subject. Participant teams of two completed a series of common visualization tasks on a set of visualizations in one of the two modalities (desktop or AR), before repeating the same set of tasks in the other modality. A mix of objective and subjective measures were recorded.

4.1.1 Study Space. The experiment was conducted in a quiet, well-lit room with a cleared 15 × 15 ft "arena" for the AR trials and a desk set up to the side for the desktop trials. Visualizations were centered in the arena and tape markings ensured participants had the same starting point for each AR trial.

4.1.2 Desktop Modality. Visualizations were presented on a 24-inch, 60Hz monitor with 1920×1080 resolution. Participants sat side-by-side and shared a single mouse and keyboard to rotate, zoom, and freely move the full-screen viewport around the scene, but could not interact with the visualization in any way. This setup (see Figure 4) and navigational model were chosen to mimic common 3D modeling and CADD workstation configurations in industry.

4.1.3 AR Modality. Participants were each wore an HMD (Microsoft HoloLens 2). The HoloLens 2 features a 60Hz 2k resolution 3:2 display for each eye and a 52 degree (measured diagonally) field

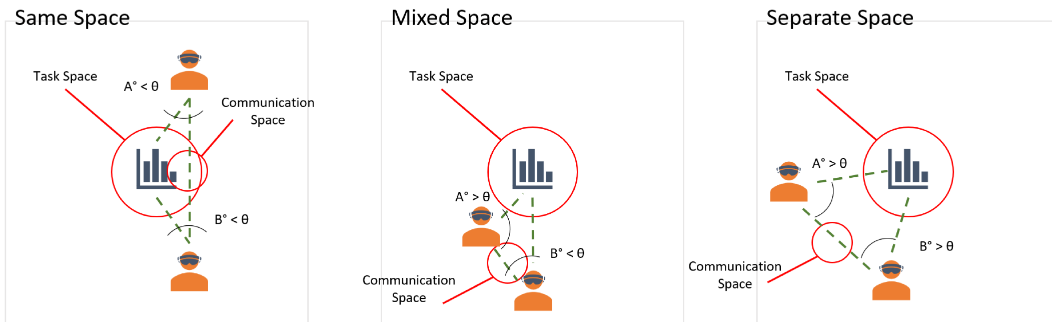

Figure 3: The three possible positional encodings. θ is the threshold angle.

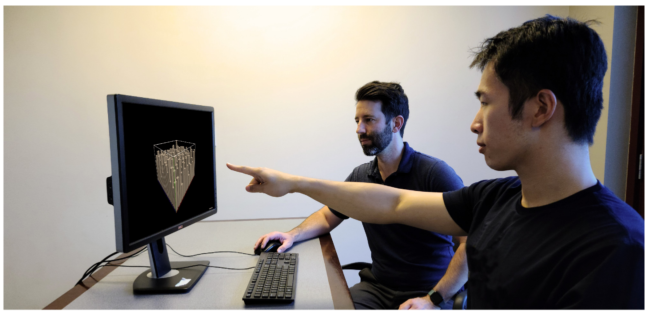

Figure 4: An example of the setup for the desktop modality.

of view [35]. Visualizations filled a $1m^3$ cube anchored to the center of the study area. Participants could freely walk around the study space but could not interact with the visualization in any way.

4.1.4 Software. Visualizations were shown in both modalities via a custom Universal Windows Platform (UWP) app for a consistent experience. The app was developed in Unity 2019 and Visual Studio 2019 using the Microsoft Mixed Reality Toolkit (MRTK) and scripted in C#. The full codebase is publicly available at https://github.com/svl-at-asu/Exploring-AR-Impact-Paper-Scripts.

4.2 Trials

Each team conducted 24 trials (two trials per modality-visualization-task combination, where $2 \times 3 \times 2 \times 2 = 24$) with the order randomized. Two trials per combination were selected to maximize the number of trials conducted, while keeping the duration of a participant's engagement with the study reasonable (about one hour total). Ten teams were recruited for the study, resulting in 240 total trials.

Due to hardware failures, seven desktop trials for one team and three AR trials for another were discarded, reducing the total to 230 trials for testing the hypotheses. Trials with a missing counterpart (same task and chart type) in the other modality were also discarded to preserve the paired test property for the statistical analysis. The resulting degrees of freedom are reported with each test in §5.

Participants were given a single task each trial. The same set of tasks was repeated in each modality, and a given task was always paired with a specific chart type. We test two types of visual analysis tasks: open and closed. Open tasks are open-ended and do not have a single correct answer. We use guided search as the basis for the open tasks, as it has been previously used to encourage discourse and communication [40]. In contrast, closed tasks have a single, correct answer. The tasks, shown in Table 1, were chosen because they are common tasks in Brehmer and Munzner's multi-level visualization taxonomy [6], and include *discovering* (or *generating*) insights by *identifying*, *comparing*, and *summarizing* visualizations, via *explore*, *browse*, and *locate* actions; notably, this taxonomy provides explicit accounting for open and closed tasks.

4.2.1 Visualizations. A mix of three chart types (data visualization techniques) commonly used for showing data in 3D [3, 8, 16, 41, 54] were used to increase the generalizability of the results. These techniques cover the primary (non-interactive) groupings for the functions of visualizations in decision-visualization environments [25]. Figure 1 shows examples of these techniques. The dataset for each of the 24 visualizations was synthetically generated using the techniques from Whitlock et al. [54] and Watts et al. [53]. Our script files publicly available at the previously-linked GitHub repository.

Table 1: This table shows the tasks given to participants for each of the task types and chart types included in the study.

	Closed Task	Open Task
Scatter Plot	Which point is the largest?	Which of the three axes (directions representing a variable) do the points suggest is the most important?
	Is there a trend? If so, what direction is it going in?	How many sub-groups (also known as clusters) would you divide these points into?
Bar Chart	Which bar is the tallest?	Find a repeating pattern in the heights of the bars.
	Is there a trend? If so, what direction is it going in?	How many categories would you divide all the bars into based on height?
Network Diagram	What is the diameter of the graph (the longest connected chain of nodes you can find, without repeating any nodes)?	How many sub-groups (also known as clusters) would you divide these nodes into?
	Which node has the most connections?	Which node is the most important?

4.3 Procedure

Each participant independently took the pre-study survey. The administrator then described the visualization techniques. Before beginning AR trials, the administrator walked participants through using the HoloLens's ob-board eyes calibration. Similarly, the desktop controls were described prior to the desktop trials.

For each trial, the administrator read participants a pre-selected question based on the randomized order from the task questions in Table 1 and then answered any questions if participants were confused about the current task or the correct procedure for using the current modality before starting the trial time. Time was not paused while participants asked such questions during the trial. Participants could freely navigate around the visualization and communicate with each other for the duration of the trial.

Trials were limited to two minutes, or when participants indicated they had an answer (whichever happened first). After each trial, participants explained their answer to the study administrator and returned to their starting positions (or reset the view on the desktop). This process was repeated for each of the 24 trials; participants completed 12 trials in one modality before switching to the other. An opportunity for a break was given between each trial.

After completing all trials, participants took a post-study survey to assess their experience and preference with the two modalities.

4.4 Data Collection

Our data analysis primarily relies on classification and coding of user behavior, which falls under the rarest category ("qualitative analysis") in Dunser et al.'s AR evaluation survey [12]. To generate a rich set of data for exploratory analysis as suggested by Lam et al. [30], we combine software logging of participant positioning with qualitative encoding from a video recording of the trials and with surveying participants via pre- and post-study questionnaires. Additionally, the completion time for each trial was recorded for comparison and for normalizing other counted measures by time.

4.4.1 Software Logging. The viewport position of each user was directly recorded by the software in six degrees of freedom sampled at 60 Hz, similarly to the study performed by Büschel et al. [8].

4.4.2 Qualitative Video Encoding. As is common in AR user studies [22, 29], trials were video and audio recorded. Three types of participant "events" were encoded: (1) gestures, (2) looks, and (3) utterances. Each recorded event was time stamped and tagged with the event type, the participant the event was for, and the team and trial number. Additional attributes were recorded for gestures and utterances. Which participant(s) controlled the mouse, and thus viewport (the "driver"), for each desktop trial was also recorded.

(1) **Gestures events** included any hand or arm motions used to communicate with the teammate. Motions not used to communicate (e.g. adjusting the HMDs or mask) were excluded. Sequential gestures were counted by either the participant lowering their arms between gestures or by a change in the train of thought between gestures (as indicated by their verbal utterances or pauses). Additionally, the gesture target (self, other participant, chart, background) and intent (reference, description, adjust view, conversational) was classified.
(2) **Look events** included visible changes of the participant's head position to look at their teammate. A continuous look was counted once regardless of duration. Cases where both teammates looked at each other were counted as two separate looks (one per teammate). This was to track the number of times participants switched their focus between the task and communication spaces, as discussed in prior work [5, 26].
(3) **Utterances** included any verbal communication with one of the following purposes: reference, position, acknowledgement, or viewport. This is based on a simplified version of a previously used scheme [28]. Consecutive utterances were broken up along significant pauses between or expressions of complete thought. Acknowledgement utterances, as used by Kraut et al. [29], were counted separately (even if part of a single sentence). Total utterances were also counted, broken down into categories based on target: a participant speaking to the other, to themself, or to the study administrator.

To evaluate the communication efficiency, deictic phrases were counted for each utterance event. Deictic phrases are any phrase that cannot be understood without the contest in which it is spoken; phrases such as "this", "those", and "here". Two types were separately counted: person deixis and spatial (place) deixis. This covers two of the three types [13, 17] and is consistent with previously used encodings [26, 28]. The third type, time deixis, was not counted because none of the tasks involved temporal data.

4.4.3 Questionnaires. User questionnaires are valuable in evaluating, and cross-referencing with performance data for, mixed reality systems [1]. Our pre-study questionnaire collected participant demographics and familiarity with computers, AR, and their study partner. The post-study questionnaire includes the NASA Task Load Index (TLX), which is commonly used in evaluative studies (e.g., [36, 52]), and free-response questions for participants to give feedback on their experience with the devices in each modality including their preference and ease of communication.

4.5 Participant Demographics

Twenty participants took part in the study (ten teams). We recruited participants who knew each other to "*facilitate rich and smooth conversation*" [26] and to lower the risk of spreading COVID-19.

Three general demographics were collected: age, gender, and education level. Participant ages ranged from 19–56 years ($\bar{x}$ = 29, σ = 13.0). Twelve participants reported male, seven female, and one nonbinary/genderqueer. For education level, seven reported "some college," three "associates degree," five "bachelors degree," four "masters degree," and one "PhD." Additionally, four study-relevant background factors were assessed in the pre-study survey using 5-point Likert scales: (1) familiarity with computers, (2) familiarity with AR/VR headsets, (3) data visualization and/or analysis, and (4) experience working with their study participant.

5 RESULTS

All statistical analysis on the dependent variables discussed in Section 4 was performed with $\alpha = 0.05$ to determine significance. G-tests and pairwise t-tests were generally employed. The degrees of freedom, t-statistic, and p-value are reported for all tests. Counts are reported both normalized and not normalized for time, as a mix of both has been done in past studies [5, 26, 28]. The collected data is included in the supplemental materials for this paper on GitHub.

5.1 Task Completion

The task completion time, in seconds, was analyzed for influences from modality, visualization type, and task type. Although not the focus of this study, the range of completion times ($n = 230, min = 2, max = 120, \bar{x} = 82.4, \sigma = 34.4$) gave confidence in a cross-sectional sample of task difficulties. The times for 220 trials were analyzed with a paired t-test, and showed no significant change across modality (t(109)=-0.205, p=0.838, |d|=-0.0219).

Although participants were given opportunities between trials to take breaks and were instructed to inform the study administrator if they wanted to stop at any time, no participant did so. One mentioned feeling some visual fatigue while using the HoloLens, but clarified they did not want to stop and never mentioned it again. No obvious signs of fatigue were observed in any study participants.

5.2 Verbal Communication

5.2.1 Quantity. Utterances were counted in one of three possible categories: *a participant speaking to the other*, *a participant speaking to themself*, and *a participant speaking to the study administrator* then aggregated in each category for each modality (desktop, AR). A Goodness of Fit (G-Test) revealed the proportions of each category varied significantly across modality (p=5.94×10^{-10}). The greatest change was in the *speaking to the study administrator* category, and is discussed more in Section 6. Rerunning the test revealed the proportion between the first two categories did not change significantly (p=0.0545). We thus conclude that the modality did not impact the amount of verbal communication, and reject **H3**.

5.2.2 Purpose. Utterance events were aggregated in each modality (desktop, AR) by purpose categories (reference, position, acknowledgement, viewport). A Goodness of Fit (G-Test) revealed the proportions of each category varied significantly across modality (p=5.10×10^{-28}). The greatest change was in the *viewport* category (198 for Desktop vs 52 for AR). This is likely explained by the fact each teammate had their own, independent viewport in AR whereas they shared a common viewport on the desktop. More communication about the viewport was necessary in the latter, as only one participant at a time could navigate the visualization.

5.2.3 Deixis. The total count of deictic phrases was made for each trial. The counts for 220 trials were analyzed with a paired t-test, and showed a small, but significant change across modality (t(109)=-2.57, p=0.0116, |d|=0.286). The same test was run on the spatial deictic phrases, and showed a similar change across modality (t(109)=2.04, p=0.0438, |d|=0.252). When normalized for time (deictic phrases per minute), the significant change in deictic phrases across modality disappeared (t(109)=-1.52, p=0.131, |d|=0.160), but the significant change in spatial deictic phrases persisted and increased slightly in effect size (t(109)=3.18, p=0.00193, |d|=0.356).

5.3 Non-verbal Communication

5.3.1 Quantity. The total count of gesture events was made for each trial. The counts for 220 trials were analyzed with a paired t-test, and showed a large and significant change across modality

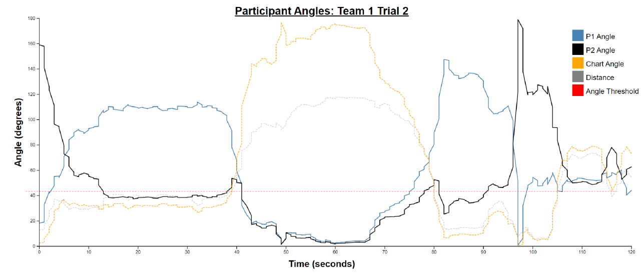

Figure 5: Example Participant Angles and Distance Over Time

(t(109)=8.96, p=1.02×10^{-14}, |d|=1.01). When normalized for time (gesture events per minute), the significant change across modality persisted and increased in effect size (t(109)=10.1, p=2.38×10^{-17}, |d|=1.24). Both changes showed an increase in gesturing in AR compared to desktop; we thus fail to reject **H1**.

A paired t-test on the 180 trials where participants gestured comparing the percentage contribution of each participant to the total gestures showed no significant change across modality (t(89)=-0.632, p=0.529, |d|=0.0898). This suggests the presence of the mouse may have increased the total gestures, as the frequent use of the mouse for gesturing was not counted.

5.3.2 Targets and Intent. Gesture events were aggregated in each modality (desktop, AR) by target categorization (self, other participant, chart, background). A Goodness of Fit (G-Test) revealed the proportions of each category varied significantly across modality (p=5.61×10^{-17}). The same test was repeated for the intent categorization (reference, description, adjust view, conversational) and similarly found a significant change (p=3.58×10^{-44}).

5.4 Overlapping Spaces

A simple exponential smoothing pass was applied to the raw location data logged by each AR headset. This data was then sampled at a rate of 2 samples per second. Because the location data depends on participant movement, a time resolution much higher than this is likely to introduce meaningless noise instead of capture intentional movement. The angles for each participant as shown in Figure 3 were calculated, along with the distance between the participants for each sample. An example of the output for one of the trials can be seen in Figure 5. This process was repeated for each AR trial.

This output is then encoded into the three positional arrangements in our proposed positional encoding method described in Section 3.3: *Same Space*, *Mixed Space*, and *Separate Space*. Figure 6 shows a visualized example from one trial. The positional data is graphed over time (left,) while the participant locations are graphed relative to the study space in a top-down view (right). The three highlighted colors each show one of the positional encodings. *Same Space* (blue) corresponds to participants on opposing sides of the visualization, (both participant angles below the threshold). *Mixed Space* (green) corresponds to participants on the same side of the visualization where one participant is between the other and the visualization (one angle above and one angle below the threshold). *Separate Space* (red) corresponds to participants on the same or adjacent sides of the visualization (both angles above the threshold).

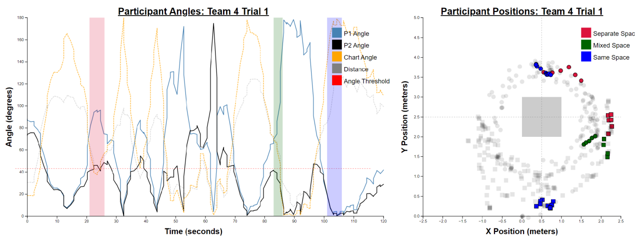

Figure 6: The connection between participant angles and positional arrangements. Colored ranges (left) and corresponding positions (right), showing the three arrangements.

The sampled positions at each time stamp were categorized into positional arrangements using a threshold of 43.3 degrees, as discussed in Section 3. The percentage of samples classified into each arrangement was then calculated for each trial, and the average percentage was calculated across all trials. Overall, teams spent an average of 34.7% time in *Same Space*, 28.2% time in *Mixed Space*, and 37.1% time in *Separate Space*. The most time was spend in the *Separate Space* arrangement, so we reject **H5**.

5.4.1 Looks. To assess how participants were shifting their focus between the communication and task spaces, the total look events for each trial was counted. The counts for 220 trials were analyzed with a paired t-test and showed a significant change across modality (t(109)=10.1, p=2.15×10^{-17}, |d|=1.27). When normalized for time (look events per minute), the change across modality persisted and increased in effect size (t(109)=11.0, p=2.45×10^{-19}, |d|=1.42).

5.4.2 Communication and Positioning. For each trial, totals of each participant event (gestures, looks, and utterances) were counted for each positional arrangement and then normalized for the time the team spent in that arrangement. An single-factor ANOVA was then used to compare the effect of positional arrangement on the resulting frequencies (events per minute) of each type of event.

Participants gestured significantly more in *Separate Space* (F(2, 222)=3.54, p=0.0308) than either of the other two spaces (t(74)=2.09, p=0.0404, |d|=0.295 compared to *Mixed* and t(74)=2.74, p=0.00756 |d|=0.420 compared to *Same*). We thus reject **H2**.

Participants looked at each other significantly more in *Same Space* (F(2, 222)=8.31, p=0.000331) than either of the other two spaces (t(74)=-4.48, p=2.70×10^{-5}, |d|=-0.672 compared to *Mixed* and t(74)=-2.42, p=0.0180, |d|=-0.364 compared to *Separate*).

Participants used significantly fewer utterances in *Same Space* (F(2, 222)=6.59, p=0.00166) than either of the other two spaces (t(74)=2.74, p=0.00774, |d|=0.455 compared to *Mixed* and t(74)=4.12, p=9.75×10^{-5}, |d|=0.583 compared to *Same*). We thus reject **H4**.

5.5 Subjective Measures

The feedback from the post-study questionnaire provides some insight into how participants perceived working in each modality.

5.5.1 NASA TLX. Participants rated their experience with the desktop computer and the AR headset on a scale of 1 to 7 in the six measures of the NASA TLX. Table 2 shows a summary of the t-tests run on their responses for each measure compared across modality.

Table 2: Pairwise T-Tests for NASA TLX Survey Responses Compared Across Modality

TLX Measure	Modality	Mean (St. Dev)	t-stat	p-value
Mental Demand	Desktop AR	4.85 (1.93) 3.45 (1.28)	t(19)= 4.08	0.000642
Physical Demand	Desktop AR	2.30 (1.66) 3.30 (1.42)	t(19)= -2.94	0.00843
Temporal Demand	Desktop AR	4.10 (1.74) 3.45 (1.23)	t(19)= 2.04	0.0554
Performance	Desktop AR	3.75 (1.77) 3.60 (1.90)	t(19)= 0.304	0.764
Effort	Desktop AR	5.30 (1.66) 3.40 (1.31)	t(19)= 5.87	0.0000182
Frustration	Desktop AR	4.65 (2.30) 2.45 (1.61)	t(19)= 4.95	0.0000898

The increased physical demand for AR is expected, as participants walked around. The decreased mental demand, effort, and frustration in AR are all consistent with subjective observations.

5.5.2 Device Preference. An overwhelming 19 of 20 (95%) participants preferred the HoloLens to the desktop. 12 participants (60%) explicitly mentioned navigation or manipulating the view in their response. Five mentioned the HoloLens being more "fun" or "engaging". The participant who didn't indicate always preferring HoloLens said they preferred the Desktop for the network diagram tasks, saying it was easier to coordinate with their teammate.

Feedback on ease of communication was mixed. 12 participants (60%) said communication was easier on desktop, five (25%) said HoloLens, and three (15%) said both equally. Participants who chose the desktop explicitly mentioned ease of "pointing" and struggled to verify they were referencing the same point in space as their teammate in AR. This suggests that pointers visible to both teammates is critical to the success and ease of users' collaboration efforts.

6 DISCUSSION AND CONCLUSION

The observed increase in gestures in AR (failing to reject **H1**), lack of change in utterances across modality (rejecting **H3**), and participants spending less than half of their time with overlapping task and communication spaces in AR (rejecting **H5**) suggest different strategies for establishing a shared understanding between collaborators are useful in AR compared to desktop.

Several factors support generalizing the results in this study. The participant demographics show a range of ages, genders, education levels, and familiarity with computers and AR headsets. We also used a mix of three common visualization types and two primary visual analysis task types. Additionally, despite the constrained study size, many of our findings show strong statistical power.

6.1 Interface Design and Navigation

Differences in device preference from the NASA TLX could be attributed to the different navigation paradigms between the desktop and AR. The "natural" navigation afforded by AR HMD devices such as the HoloLens lessens the gulfs of evaluation and execution [20], and may explain the reported lower frustration and demanded effort in AR. Our choice of navigation paradigm represents common setups in industry but may limit generalizing some of the comparative results between desktop and AR. Other desktop paradigms (such as direct manipulation) should be compared in future work.

6.2 Communication

At first, the increased reliance on gestures in AR suggests that participants were more confident their teammate shared the same view of the task space when wearing the AR HMDs than when both looked at the same desktop monitor. However, the small change in verbal communication (utterances) in AR suggests that the increased gestures were not taking its place; teammates communicated more overall in AR by augmenting their verbal communication with gestures. This is further suggested by the increase in spatial deixis observed in AR. This, combined with the lack of increase in utterances (even when normalized for trial time) suggests the communication efficiency (deixis per amount of utterances) was higher in AR.

While not contradicting Kiyokawa et al. [26], these findings do question the reasoning behind how AR impacts the necessity of verbal and nonverbal communication between collaborators. Further research on the content and purpose of both utterances and gestures is needed, perhaps using the Inter-referential Life Cycle model used by Chastine et al. [9] or a similar method used by Kraut et al. [29]. The role each plays in teammates' efforts to establish conversational grounding may change between desktop and AR.

6.3 Positional Encoding

Participants spent the most time in the *Same Space* arrangement and the least in *Mixed Space* on average across all trials, and communicated differently in each. Participants gestured and verbally communicated more in *Separate Space* but looked at each other more in *Same Space*. This suggests both arrangements, which interestingly correspond to ones illustrated by Billinghurst et al. [5], play different roles in collaboration efforts. Our positional coding method provides quantitative evidence consistent with prior observations of participants moving to see the "exact view" of their teammate to communicate better in AR [9], as well as reducing 3D tasks to 2D tasks while performing visual analysis in VR [8]. Participants moving to and communicating more while in the *Same Space* arrangement suggest similar strategies in our study.

A subjective review of the trial recordings supports this. *Same Space* was often present when teammates were discussing the visualization as a whole or broad strategy or independently exploring while *Separate Space* was often present while teammates referenced specific parts of the visualizations or "synchronized" their views. More work should be done on assessing the role these arrangements play in facilitating different collaborative behaviors.

An interface implementing an overview-plus-detail design pattern [44] by mixing linked 2D and 3D visualizations (a 3D "overview" to provide context and allow participants to select and pop out detailed 2D views from) may be useful in supporting these behaviors. Making such a 2D detail view visible to all teammates would satisfy some of the issues participants raised when discussing device preferences (Section 5.5.2) that frustrated establishing conversational grounding between teammates, however how such views can be integrated for both participants in an intuitive and helpful manner within an AR context remains an open research question.

The novel method for encoding positional arrangements outlined in Section 3.3 provided a strong quantitative basis for identifying these positional arrangements and correlating other encoded data with them, and should assist in increasing the replicability of studies that use it. The method also provides opportunities for powerful data visualizations to assist with the analysis of the associated data, such as the one presented in Figure 6. Developing visualization tools to assist researchers in quickly encoding the positional arrangements based on this technique would be a good pursuit for future work, as would expanding this technique to larger teams.

6.4 Study Limitations

Our proposed positional encoding method relies on three major assumptions. First, we assume that all participants are viewing the same, public scene of virtual objects; if an object exists for one user, it also exists for the others, anchored in the same position in the same space relative to the physical world. This is consistent with prior work [5, 26], but still leaves out systems where users can independently control private views of the visualization. Extending our method to include such systems is left to future work.

Second, extrapolating participants' views of their collaborator and the visualization assumes participants faced nearly centered on their calculated angle. A participant could, for example, face away from both and thus be able to see neither. While a subjective review of the trial recordings suggests this is a reasonable assumption (participants generally faced either inwards towards the chart or towards their teammate), combining more robust gaze and head orientation tracking with our method is left to future work.

Third, we assume the interactive interfaces are part of the same task space as the visualization itself. This assumption is made to consider "seams" between the task and communication spaces consistent with prior work [4, 23], but leaves out the possibility of such "seams" existing between a decoupled visualization and interaction tool (such as a remote or controller). Our study did not include interaction; thus this assumption held. Exploring the validity of generalizing this assumption is also left to future work.

Our user study is small (20 participants), but nearly double the median size (12) of previously surveyed collaborative AR studies [10]. As mentioned in §4.5, our demographics gave confidence in the representative nature of the study, and we chose a within-subject study design (see §4.1) to increase the statistical power of our results (consistent with the overwhelming majority of previous collaborative AR studies [10]). Despite this, validating the generalizability of our results would require a significantly larger study. Additionally, future studies can expand on the types of tasks that are tested.

6.5 Conclusion

We present a novel method for encoding the positional arrangement of pairs of co-located, synchronous collaborators using AR HMDs. Our method adapts a previously-used method for the less-constrained participant movements afforded by HMDs and increases its study replicability by basing the encoding off of quantitative data. We also demonstrated our method's use in evaluating collaborative visualization behaviors through a user study of collaborating dyads in AR. The results challenge prior assumptions about the role positioning plays in AR and show that our proposed encoding method can help evaluate, visualize, and analyze this role.

ACKNOWLEDGMENTS

This research was supported in part by the U.S. National Science Foundation through grant DUE-2216542. The authors would also like to thank the Decision Theater at Arizona State University for their support in conducting the study.

REFERENCES

[1] Cédric Bach and Dominique L Scapin. 2004. Obstacles and perspectives for evaluating mixed reality systems usability. In *Acte du Workshop MIXER, IUI-CADUI*, Vol. 4. Citeseer.

[2] Kevin M Baird and Woodrow Barfield. 1999. Evaluating the effectiveness of augmented reality displays for a manual assembly task. *Virtual Reality* 4, 4 (1999), 250–259.

[3] Daniel Belcher, Mark Billinghurst, SE Hayes, and Randy Stiles. 2003. Using augmented reality for visualizing complex graphs in three dimensions. In *The Second IEEE and ACM International Symposium on Mixed and Augmented Reality, 2003. Proceedings.* IEEE, 84–93.

[4] Mark Billinghurst and Hirokazu Kato. 1999. Collaborative mixed reality. In *Proceedings of the First International Symposium on Mixed Reality*. 261–284.

[5] Mark Billinghurst, Hirokazu Kato, Kiyoshi Kiyokawa, Daniel Belcher, and Ivan Poupyrev. 2002. Experiments with face-to-face collaborative AR interfaces. *Virtual Reality* 6, 3 (2002), 107–121.

[6] Matthew Brehmer and Tamara Munzner. 2013. A multi-level typology of abstract visualization tasks. *IEEE transactions on visualization and computer graphics* 19, 12 (2013), 2376–2385.

[7] Wolfgang Broll, Irma Lindt, Jan Ohlenburg, Michael Wittkämper, Chunrong Yuan, Thomas Novotny, Chiron Mottram, Andreas Strothmann, et al. 2004. Arthur: A collaborative augmented environment for architectural design and urban planning. *JVRB-Journal of Virtual Reality and Broadcasting* 1, 1 (2004).

[8] Wolfgang Büschel, Patrick Reipschläger, Ricardo Langner, and Raimund Dachselt. 2017. Investigating the use of spatial interaction for 3d data visualization on mobile devices. In *Proceedings of the 2017 ACM International Conference on Interactive Surfaces and Spaces*. 62–71.

[9] Jeffrey W Chastine, Kristine Nagel, Ying Zhu, and Luca Yearsovich. 2007. Understanding the design space of referencing in collaborative augmented reality environments. In *Proceedings of graphics interface 2007*. 207–214.

[10] Arindam Dey, Mark Billinghurst, Robert W Lindeman, and J Swan. 2018. A systematic review of 10 years of augmented reality usability studies: 2005 to 2014. *Frontiers in Robotics and AI* 5 (2018), 37.

[11] Suyang Dong, Amir H Behzadan, Feng Chen, and Vineet R Kamat. 2013. Collaborative visualization of engineering processes using tabletop augmented reality. *Advances in Engineering Software* 55 (2013), 45–55.

[12] Andreas Dünser, Raphaël Grasset, and Mark Billinghurst. 2008. *A survey of evaluation techniques used in augmented reality studies.* Human Interface Technology Laboratory New Zealand.

[13] Ardita Dylgjeri and Ledia Kazazi. 2013. Deixis in modern linguistics and outside. *Academic Journal of Interdisciplinary Studies* 2, 4 (2013), 87–87.

[14] Rob Edsall and Kelli L Larson. 2006. Decision making in a virtual environment: Effectiveness of a semi-immersive "decision theater" in understanding and assessing human-environment interactions. In *Proceedings of AutoCarto*, Vol. 6. 25–28.

[15] Clarence A Ellis, Simon J Gibbs, and Gail Rein. 1991. Groupware: some issues and experiences. *Commun. ACM* 34, 1 (1991), 39–58.

[16] Niklas Elmqvist, Pierre Dragicevic, and Jean-Daniel Fekete. 2008. Rolling the dice: Multidimensional visual exploration using scatterplot matrix navigation. *IEEE transactions on Visualization and Computer Graphics* 14, 6 (2008), 1539–1148.

[17] Charles J Fillmore. 1966. Deictic categories in the semantics of 'come'. *Foundations of language* (1966), 219–227.

[18] Susan R. Fussell, Leslie D. Setlock, and Elizabeth M. Parker. 2003. Where Do Helpers Look? Gaze Targets during Collaborative Physical Tasks. In *CHI '03 Extended Abstracts on Human Factors in Computing Systems* (Ft. Lauderdale, Florida, USA) *(CHI EA '03)*. Association for Computing Machinery, New York, NY, USA, 768–769. https://doi.org/10.1145/765891.765980

[19] Darren Gergle, Robert E Kraut, and Susan R Fussell. 2004. Action as language in a shared visual space. In *Proceedings of the 2004 ACM conference on Computer supported cooperative work*. 487–496.

[20] Edwin L Hutchins, James D Hollan, and Donald A Norman. 1985. Direct manipulation interfaces. *Human–computer interaction* 1, 4 (1985), 311–338.

[21] Kori Marie Inkpen. 1997. *Adapting the human-computer interface to support collaborative learning environments for children.* Ph. D. Dissertation. University of British Columbia.

[22] Petra Isenberg, Danyel Fisher, Sharoda A Paul, Meredith Ringel Morris, Kori Inkpen, and Mary Czerwinski. 2011. Co-located collaborative visual analytics around a tabletop display. *IEEE Transactions on visualization and Computer Graphics* 18, 5 (2011), 689–702.

[23] Hiroshi Ishii, Minoru Kobayashi, and Kazuho Arita. 1994. Iterative design of seamless collaboration media. *Commun. ACM* 37, 8 (1994), 83–97.

[24] Robert Johansen, David Sibbet, Suzyan Benson, Alixia Martin, Robert Mittman, and Paul Saffo. 1991. *Leading business teams: How teams can use technology and group process tools to enhance performance.* Addison-Wesley Longman Publishing Co., Inc.

[25] Beatrice John, D Lang, Henrik von Wehrden, Ruediger John, and Arnim Wiek. 2018. Advancing decision-visualization environments: Empirically informed design guidelines.

[26] Kiyoshi Kiyokawa, Mark Billinghurst, Sohan E Hayes, Anoop Gupta, Yuki Sannohe, and Hirokazu Kato. 2002. Communication behaviors of co-located users in collaborative AR interfaces. In *Proceedings. International Symposium on Mixed and Augmented Reality*. IEEE, 139–148.

[27] Robert E Kraut, Susan R Fussell, and Jane Siegel. 2003. Visual information as a conversational resource in collaborative physical tasks. *Human–computer interaction* 18, 1-2 (2003), 13–49.

[28] Robert E Kraut, Darren Gergle, and Susan R Fussell. 2002. The use of visual information in shared visual spaces: Informing the development of virtual co-presence. In *Proceedings of the 2002 ACM conference on Computer supported cooperative work*. 31–40.

[29] Robert E Kraut, Mark D Miller, and Jane Siegel. 1996. Collaboration in performance of physical tasks: Effects on outcomes and communication. In *Proceedings of the 1996 ACM conference on Computer supported cooperative work*. 57–66.

[30] Heidi Lam, Enrico Bertini, Petra Isenberg, Catherine Plaisant, and Sheelagh Carpendale. 2011. Empirical studies in information visualization: Seven scenarios. *IEEE transactions on visualization and computer graphics* 18, 9 (2011), 1520–1536.

[31] Matthew Lombard and Theresa Ditton. 1997. At the heart of it all: The concept of presence. *Journal of computer-mediated communication* 3, 2 (1997), JCMC 321.

[32] Tim Marsh. 2003. Staying there: an activity-based approach to narrative design and evaluation as an antidote to virtual corpsing. In *Being there: Concepts, effects and measurement of user presence in synthetic environments, studies on new technologies and practices in communication series.* Citeseer.

[33] Andreas Mathisen, Tom Horak, Clemens Nylandsted Klokmose, Kaj Grønbæk, and Niklas Elmqvist. 2019. InsideInsights: Integrating Data-Driven Reporting in Collaborative Visual Analytics. In *Computer Graphics Forum*, Vol. 38. Wiley Online Library, 649–661.

[34] Mark McGill, Daniel Boland, Roderick Murray-Smith, and Stephen Brewster. 2015. A dose of reality: Overcoming usability challenges in vr head-mounted displays. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. 2143–2152.

[35] Microsoft. 2020. HoloLens 2-Overview, Features, and Specs: Microsoft HoloLens. https://www.microsoft.com/en-us/hololens/hardware. https://www.microsoft.com/en-us/hololens/hardware.

[36] Jens Müller, Johannes Zagermann, Jonathan Wieland, Ulrike Pfeil, and Harald Reiterer. 2019. A Qualitative Comparison Between Augmented and Virtual Reality Collaboration with Handheld Devices. In *Proceedings of Mensch und Computer 2019*. 399–410.

[37] Tamara Munzner. 2014. *Visualization analysis and design.* CRC press.

[38] Andrew YC Nee, SK Ong, George Chryssolouris, and Dimitris Mourtzis. 2012. Augmented reality applications in design and manufacturing. *CIRP annals* 61, 2 (2012), 657–679.

[39] Jiazhi Ou, Lui Min Oh, Jie Yang, and Susan R. Fussell. 2005. Effects of Task Properties, Partner Actions, and Message Content on Eye Gaze Patterns in a Collaborative Task. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Portland, Oregon, USA) *(CHI '05)*. Association for Computing Machinery, New York, NY, USA, 231–240. https://doi.org/10.1145/1054972.1055005

[40] Kyoung S Park, Abhinav Kapoor, and Jason Leigh. 2000. Lessons learned from employing multiple perspectives in a collaborative virtual environment for visualizing scientific data. In *Proceedings of the Third International Conference on Collaborative Virtual Environments*. 73–82.

[41] Dheva Raja, Doug Bowman, John Lucas, and Chris North. 2004. Exploring the benefits of immersion in abstract information visualization. In *Proc. Immersive Projection Technology Workshop*. 61–69.

[42] Dieter Schmalstieg, Anton Fuhrmann, Gerd Hesina, Zsolt Szalavári, L Miguel Encarnaçao, Michael Gervautz, and Werner Purgathofer. 2002. The studierstube augmented reality project. *Presence: Teleoperators & Virtual Environments* 11, 1 (2002), 33–54.

[43] Mickael Sereno, Xiyao Wang, Lonni Besancon, Michael J Mcguffin, and Tobias Isenberg. 2020. Collaborative Work in Augmented Reality: A Survey. *IEEE Transactions on Visualization and Computer Graphics* (2020), 1–1. https://doi.org/10.1109/TVCG.2020.3032761

[44] Ben Shneiderman. 2003. The eyes have it: A task by data type taxonomy for information visualizations. In *The craft of information visualization*. Elsevier, 364–371.

[45] Rebecca R Springmeyer, Meera M Blattner, and Nelson L Max. 1992. A characterization of the scientific data analysis process. In *Proceedings of the 3rd conference on Visualization'92*. IEEE Computer Society Press, 235–242.

[46] Zsolt Szalavári, Dieter Schmalstieg, Anton Fuhrmann, and Michael Gervautz. 1998. "Studierstube": An environment for collaboration in augmented reality. *Virtual Reality* 3, 1 (1998), 37–48.

[47] Arthur Tang, Frank Biocca, and Lynette Lim. 2004. Comparing differences in presence during social interaction in augmented reality versus virtual reality environments: An exploratory study. *Proceedings of PRESENCE* (2004), 204–208.

[48] Arthur Tang, Charles Owen, Frank Biocca, and Weimin Mou. 2003. Comparative effectiveness of augmented reality in object assembly. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. 73–80.

[49] Anthony Tang, Melanie Tory, Barry Po, Petra Neumann, and Sheelagh Carpendale. 2006. Collaborative coupling over tabletop displays. In *Proceedings of the SIGCHI conference on Human Factors in computing systems*. 1181–1190.

[50] Melanie Tory, Arthur E Kirkpatrick, M Stella Atkins, and Torsten Moller. 2005. Visualization task performance with 2D, 3D, and combination displays. *IEEE transactions on visualization and computer graphics* 12, 1 (2005), 2–13.

[51] April Yi Wang, Anant Mittal, Christopher Brooks, and Steve Oney. 2019. How data scientists use computational notebooks for real-time collaboration. *Proceedings of the ACM on Human-Computer Interaction* 3, CSCW (2019), 1–30.

[52] Xiangyu Wang and Phillip S Dunston. 2011. Comparative effectiveness of mixed reality-based virtual environments in collaborative design. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)* 41, 3 (2011), 284–296.

[53] Duncan J Watts and Steven H Strogatz. 1998. Collective dynamics of 'small-world' networks. *nature* 393, 6684 (1998), 440–442.

[54] Matt Whitlock, Stephen Smart, and Danielle Albers Szafir. 2020. Graphical Perception for Immersive Analytics. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE, 616–625.

[55] Jian Zhao, Michael Glueck, Petra Isenberg, Fanny Chevalier, and Azam Khan. 2017. Supporting handoff in asynchronous collaborative sensemaking using knowledge-transfer graphs. *IEEE transactions on visualization and computer graphics* 24, 1 (2017), 340–350.