Cone of Vision as a Behavioural Cue for VR Collaboration

RICCARDO BOVO, Imperial College London, United Kingdom
DANIELE GIUNCHI, University College London, United Kingdom
MUNA ALEBRI, University College London, United Kingdom
ANTHONY STEED, University College London, United Kingdom
ENRICO COSTANZA, University College London, United Kingdom
THOMAS HEINIS, Imperial College London, United Kingdom

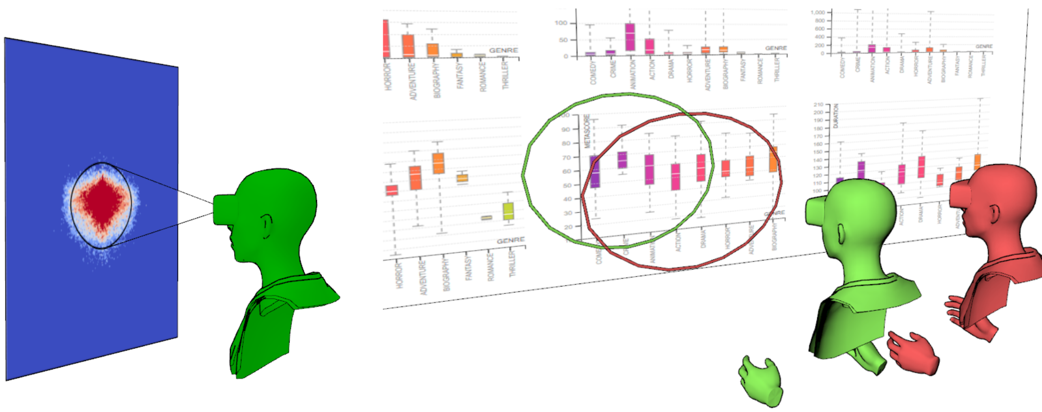


Fig. 1. Left image: cone of vision modelled with average fixation map, that is an updated version of the field of view frustum visualisation. Right Image: bi-directional attention cues between two users that are collaborating on an EDA task.

Mutual awareness of visual attention is essential for collaborative work. In the field of collaborative virtual environments (CVE), it has been proposed to use Field-of-View (FoV) frustum visualisations as a cue to support mutual awareness during collaboration. Recent studies on FoV frustum visualisations focus on asymmetric collaboration with AR/VR hardware setups and 3D reconstructed environments. In contrast, we focus on the general-purpose CVEs (i.e., VR shared offices), whose popularity is increasing due to the availability of low-cost headsets, and the restrictions imposed by the pandemic. In these CVEs collaboration roles are symmetrical, and the same 2D content available on desktop computers is displayed on 2D surfaces in a 3D space (VR screens). We prototyped one such CVE to evaluate FoV frustrum visualisation within this collaboration scenario. We also implement a FoV visualisation generated from an average fixation map (AFM), therefore directly generated by users' gaze behaviour which we call Cone of Vision (CoV). Our approach to displaying the frustum visualisations is tailored for 2D surfaces in 3D space and allows for self-awareness of this visual cue. We evaluate CoV in the context of a general exploratory data analysis (EDA) with 10 pairs of participants. Our findings indicate that CoV is beneficial during shifts between independent and collaborative work and supports collaborative progression across the visualisation. Self-perception of the CoV improves

Authors' addresses: Riccardo Bovo, Imperial College London, London, United Kingdom, rb1619@ic.ac.uk; Daniele Giunchi, University College London, London, United Kingdom, d.giunchi@ucl.ac.uk; Muna Alebri, University College London, London, United Kingdom, m.alebri@ucl.ac.uk; Anthony Steed, University College London, London, United Kingdom, a.steed@ucl.ac.uk; Enrico Costanza, University College London, London, United Kingdom, e.costanza@ucl.ac.uk; Thomas Heinis, Imperial College London, London, United Kingdom, t.heinis@ic.ac.uk.

visual attention coupling, reduces the number of times users watch the collaborator's avatars and offers a consistent representation of the shared reality.

CCS Concepts: • **Human-centered computing → Computer supported cooperative work**; **Empirical studies in HCI**; **Virtual reality**; **Visual analytics**.

Additional Key Words and Phrases: Field of View frustum visualizations, visual attention cues, VR collaborative analytics

ACM Reference Format:
Riccardo Bovo, Daniele Giunchi, Muna Alebri, Anthony Steed, Enrico Costanza, and Thomas Heinis. 2022. Cone of Vision as a Behavioural Cue for VR Collaboration. *Proc. ACM Hum.-Comput. Interact.* 6, CSCW2, Article 502 (November 2022), 27 pages. https://doi.org/10.1145/3555615

1 INTRODUCTION

While immersive collaborative virtual environments (ICVE) are often tailored for highly specialised tasks and specific hardware setups, a new range of VR productivity software that caters for virtual office meetings, presentations and general data collaborative applications is becoming available [4, 5, 21, 36, 55]. These applications allow the display of 2D desktop content (i.e., spreadsheets, webpages, etc.) as 2D surfaces in the virtual environment enabling users to run collaborative meetings in three-dimensional (3D) shared virtual offices instead of traditional video teleconference formats. Such general applications broaden the range of possible remote productivity and collaborative scenarios that can be carried out with low-cost, consumer-oriented popular virtual reality (VR) headsets (e.g., Oculus Quest). For example, Cross-virtuality Analytics (XVA) makes use of immersive and visual analytics, providing collaborative interfaces to support multiple users across the reality–virtuality continuum as outlined by Fröhler et al. [30] and explored in numerous studies [15, 16, 43, 46, 52, 59, 64, 69]. Further examples of XVA are the works of Cavallo et al. [17, 18], where visual analytics consist of a general exploratory data analysis task (EDA).Such scenarios are gaining additional interest because of the COVID pandemic, which limited in-person meetings and presentations and promoted remote working. While collaboration in an ICVE has been suggested to be as beneficial as physical co-location [35, 60, 68], it lacks some of the non-verbal communication cues (i.e. gaze awareness), which enrich in-person collaboration [39]. Mutual awareness of visual focus is an essential part of collaborative work. It refers to the ability to interpret one another's focus of visual attention. In video teleconferencing, many studies (e.g. [49, 78]) convey the importance of gaze awareness. The ability to effectively communicate one's focus of visual attention and interpret the focus of others is necessary during collaborative scenarios as it allows collaborators to coordinate their focus on a specific feature [20, 71]. Previous studies demonstrate that mutual awareness of visual focus can impact communication and performance [20, 71].

While representing eye-gaze direction would be the most obvious way to represent visual attention, most low-cost VR head-mounted displays (HMDs) do not provide eye-tracking. Moreover, eye-tracking suffers from low accuracy due to calibration errors and the drift of eye-tracking sensors [41] and the eye traits of different ethnic groups can lead to eye-trackers failures [12]. Limited accuracy in a collaborative scenario may lead to angular displacement that could convey the wrong information between participants. To support the broadest range of devices and develop a fallback method for eye-tracking failure, we focus on sharing visual attention modalities that do not rely on gaze recognition but instead use the head orientation as an approximation of gaze [6, 45] such as field of view (FoV) visualisations [34, 58, 59].

FoV frustum visualisations were proposed as behavioural cues for collaborative virtual environments (CVEs) more than 20 years ago by Hindmarsh et al. [34]. They have also been explored in the context of mixed reality collaboration, demonstrating the benefits of such visualisations in enhancing mutual awareness of visual focus among collaborators Piumsomboon et al. [58, 59]. However, those studies explored asymmetric (AR/VR) benchmark tasks consisting of cooperative search and placement of virtual objects. Such tasks are important in many collaborative situations but do not address mutual awareness of visual focus during collaboration in VR offices where collaboration is symmetric and the data displayed is 2D data (i.e., spreadsheets, web pages, presentations). To the best of our knowledge, no researches aim to evaluate FoV frustum visualisation in the context of XVA applications where the observed visualisations are displayed on 2D VR screens. We asked participants to perform a general EDA task [18] in a symmetric collaborative context (VR/VR), which is a common real-case scenario of remote working.

Furthermore, previous studies suggested that users should be free to show or hide FoV representations or that an intelligent software application should be able to understand when FoV representations should be shown automatically based on collaborative dynamics [50, 59]. However, they do not report which collaborative dynamics benefit from the use of such FoV visualisations and in which cases FoV visualisations disturb or even prevent the completion of a task. To address this gap, this paper reports on a study designed to identify in which phases of a general collaborative EDA task the FoV visualisation is most beneficial. This is so that future applications of FoV visualisations may support autonomous showing and hiding of FoV frustum visualisations depending on the phase the collaboration is in.

We propose an alternative visualisation to the FoV frustum, which we call "Cone of Vision" (CoV), specifically tailored for 2D content displayed as a 2D surface (VR screen) in the 3D environment. While the FoV represents the whole area or volume where the gaze may be (based on the head position), there are parts of the FoV where the gaze is more likely to be. Using users' gaze samples from a VR dataset by Agtzidis et al. [1], we generated a CoV that represents the area with 70% probability of containing the users' fixations, which is about $\frac{1}{9}$ the area size of the FoV frustum depiction. We visualise the CoV by depicting the intersection between the cone and the 2D surface. This design reduces concerns about the occlusion of the observed data and the size of the FoV [59]. Such a design allows users to observe their own CoV (bi-directional visual cue), which is a feature that previous research has not explored. Previous visualisation techniques were not prone to FoV self-perception, only allowing the collaborator to see the FoV (mono-directional visual cue).

We report on a remote user experiment where ten pairs of participants experienced three conditions within-group: *no-CoV*, *CoV* of the collaborator (mono-directional visual cue), and *CoV plus self-CoV* (bi-directional visual cue), as illustrated in Figure 3. In each condition, each pair of participants was asked to perform a collaborative EDA task with the goal of extracting as many correct insights from a given data visualisation. The environment layout was based on prior works [17, 26, 43, 69] as outlined in Section 2.4. Our contribution is three-fold and consists in:

- introducing a visual attention cue tailored for 2D VR screens and identifying whether visual attention cues are beneficial during collaborative EDA tasks and particularly understanding in which phases they are most useful, extending the work of Piumsomboon et al. [59];
- exploring the benefits of visual attention cues self-awareness by comparing mono-directional cues (which can be seen only by collaborators) and bi-directional cues (which can be seen by both users), extending the work of Jing et al. [39];
- exploring and identifying synergies between laser pointers and CoV extending prior work on deictic pointing in CVEs [81, 82].

Our results show that the CoV facilitate participants in shifting from independent to collaborative work. Also, CoV visualisations support collaborative progression across different visualisations and improve understanding of laser pointing actions. Consistent with previous studies [59], CoV

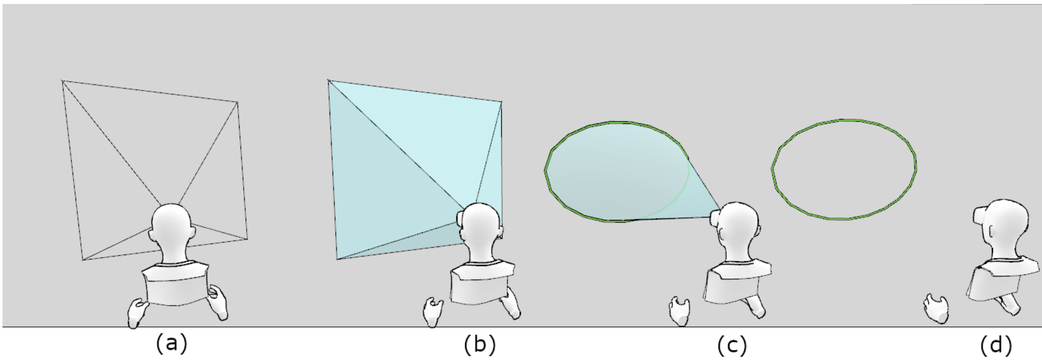


Fig. 2. (a) Hindmarsh et al. [34] propose to visualise the FoV frustum as a wireframe (first on the left). (b) Piumsomboon et al. [59] propose to visualise the FoV frustum shaded (middle). (d) Rather than displaying the FoV frustum, we propose to display the intersection of a custom calculated cone of vision (CoV) with the observed data. (c) during our pilot phase we also test a shaded version of the CoV.

visualisation improves mutual awareness of visual attention. Additionally, it enables participants to achieve the same number of correct insights with a lower cognitive effort. Most importantly, our results highlight how the novelty of self-perception of the CoV improves visual attention coupling, reduces the number of times users watch the collaborator's avatar, and ultimately offers a consistent representation of the shared environment to participants.

2 LITERATURE REVIEW

2.1 FoV Frustum Visualisations

FoV frustum visualisations have been proposed as behavioural cues for CVEs over 20 years ago, first by Hindmarsh et al. [34] and followed by Fraser et al. [29] (Figure 6). Such visualisations aid collaboration by improving mutual awareness of visual focus among pairs of collaborators [19, 59]. The graphical representation of the FoV frustum should match requirements related to the typology of visualised data and the specific tasks/interactions for such applications. The early studies represented the frustum as a wire-frame shape [29, 34]. In contrast, more recent studies explored the depiction of the frustum with semi-opaque shaders [8, 59, 61]. There are trade-offs to these representations. For example, Piumsomboon et al. [59] report problems related to the shader's opacity, which were sometimes the cause of visibility problems for users. Furthermore, the FoV size of VR headsets has also been a cause of concern Piumsomboon et al. [59]. As the FoV size increases (e.g. the Valve Index has a specified FoV of 130 degrees horizontal), a direct representation of the FoV could appear ambiguous. [59] Propose to reduce it via trial and error to a size which better represents the user's visual attention. In contrast, we use a data-informed approach, which allows us to reduce the size (as well as to adapt its shape) of the FoV visualisation based on the statistical density obtained from an average fixation map (section 3.2.3). Piumsomboon et al. [59] suggest that an intelligent system could autonomously choose when to show or hide FoV frustum visualisations, based on the detection of collaborative context. However, while suggesting such an idea, the study does not explore in which phases of collaboration the FoV visualisation should be displayed. To understand in which stages of our collaborative scenario the FoV frustum visualisation is most useful, we administer a semi-structured interview and we run a qualitative analysis by coding the interview transcripts. The outcome of our qualitative analysis allows us to have a clear idea of which stages of collaboration benefit the most from FoV visualisations (section 4.2).

2.2 Cost and Limitations of Eye-tracking on VR devices

Depicting the direction of eye-gaze would be the most obvious way to represent visual attention in collaborative settings [39]. Including eye-tracking technology in HMDs is feasible nowadays but impacts their prices considerably. For example, at the time of writing the Pico Neo 3 with eye-tracking costs 50% more than the basic version (699$ to 899$), and the HTC Vive Pro has a 100% increment in the price for the version with an eye-tracker (699$ to 1399$). Because of such costs differences, HMD devices with eye tracking are not popular in the consumer market. For example, the number of Virtual reality (VR) headset unit sales worldwide in the 4th quarter of 2019 and 4th quarter of 2020 shows that the vast majority of VR devices sold don't have eye-tracking [3]. While such a lack of popularity in the consumer market of HMDs with eye-tracking affects the potential pool size of participants for a remote VR study during the pandemic, it is plausible that eye-tracking devices will become cheaper and more popular with time. However, whether eventually all VR HMDs will have eye-tracking technologies remains questionable. Additional hardware costs are not the only reason for exploring methods alternative eye-tracking.

Eye-tracking can suffers from low accuracy due to calibration errors and the drift of wearable eye-tracking sensors [41]. While eye-tracking was used in HCI studies with different purposes such as selection [62], object manipulation [84], and collaboration [33], previous work shows that participant data is often discarded due to calibration and tracking issues [2, 7, 25]. Participant behaviour may cause the eye tracker to slip on the participant's head, potentially strongly affecting data quality [51]. Moreover, eyes traits of ethnic groups may affect the accuracy of eye trackers. For example, a study by [12] shows that Asian participants' eye-trackers accuracy was worse than that of African and Caucasian participants. The problems mentioned above affect the accuracy in such a way that, in a collaborative scenario with dense data visualisation, a small angular displacement may convey the wrong information between participants.

Indeed, from the studies mentioned above, we deduce that the HCI research community and VR developers should explore alternative methods to provide visual attention information for eye-tracker-less VR HMDs and fallback mechanisms when the eye-tracker loses accuracy or fails.

2.3 VR Screens

In the last few years, thanks to the availability of low-cost VR HMDs such as Oculus Quest, several commercial ICVEs became available [4, 5, 21, 36, 55]. These applications allow the display of standard 2D documents on VR screens in a 3D environment where multiple users can collaborate while embodied in avatars with the support of laser pointers and sketching tools. These VR applications replicate the physical offices spaces giving users the ability to load their files from local or cloud drives. The files supported consist of documents such as power points, excel, videos and web pages. However, none of these applications implements the FoV frustum visualisations, as noted by Wong and Gutwin [82]. Similar research ICVE frameworks for collaboration are published in the literature [17, 18, 24, 43, 64, 83], but only a few implement FoV frustum visualisations. These studies are tailored for specific asymmetric tasks with specific Mixed Reality (MR) hardware setups. These tasks mainly consist of collaboration in VR/AR object placement tasks in a static 3D reconstructed environment [58, 59] or a scene reconstructed in real-time by clusters of RGBD sensors [8].

Such 3D data, asymmetric collaborative scenarios, and specialised hardware do not validate the use of FoV frustum visualisations in collaborative scenarios with low-cost VR HMDs over 2D data and generic EDA tasks. As the scenario we propose (i.e., 2D content and low-cost VR headset) has a potentially broad application, we fill the gap in the literature by conducting a study to explore FoV visualisation within it. We explore the use of FoV frustum visualisation in VR offices with 2D

VR screens. Thus, we excluded from our review such XVA works where the visualisation uses the 3rd dimension developed explicitly for the VR environment [15, 16, 46].

2.4 VR Screens Layouts

Environment configuration and 2D panels positions and orientations can be arranged in different ways in an immersive virtual scenario. Such a setup may affect participants' perception and thus impacts the sense-making process as well as the collaboration. Previous works related to HMD headsets to support everyday computing tasks with desks and keyboards opted towards retaining a cylindrical display layout similar to the one present on physical working desktops and shaped accordingly to ergonomic mechanics of eyes and head [47, 56]. This cylindrical layout approach is also used by applications such as Oculus Infinite Office [48] and has been explored in AR studies such as in [47, 56]. The works of Google Daydream, when introducing the virtual screen concept, exploit the idea of screens in a cylindrical contour that gives a less claustrophobic feeling and an improved parallax effect [31].

When users are not constrained by a desk or a keyboard, for example, if they are standing and free to move in the space, the optimal layout could change. The work of Satriadi et al. [69] explores the 3D placement of 2D surfaces for individual tasks related to visualisations mapping. They study how users arrange the layout of 2D surfaces on 3D space, and their findings highlight how users tend to prefer semi-spherical and spherical layouts. Further studies by Lee et al. [43] highlight considerations related to the placement of 2D VR screen in the 3D space when collaborating with others users. They suggest that when 2D visualisations are placed for collaboration purposes, participants place them neatly on the wall of the virtual environments. This careful placement makes the 2D screen content easily visible to others. Moreover, they show that participants can present their visualisation to others when this wall layout is used. Their findings are consistent with hybrid reality environments for immersive data analysis such as Cavallo et al. [17], Febretti et al. [26] in which 2D screens define the walls/boundaries of the environments. Contrary to Lee et al. [43] which was limited by the physical spaces available to him (which we estimate to be 3m x 3m square room) the Hybrid Reality Environment for Collaborative Information Analysis used by [17, 26] are cylindrical and therefore more in line with the findings from [69]. Not being limited by physical space, we designed our virtual environment to match the same size and circular shape (3 m radius) as the hybrid reality environment [17, 26] as shown in Figure 4. Although [69] describes the layout as spherical because of the increment in size due to the collaborative context, the spherical curvature can be approximated to a cylinder.

2.5 Mutual Awareness of Visual Focus

The purpose of an FoV frustum visualisation is to support the awareness of visual focus among collaborators during collaborative work. Mutual awareness of visual focus is an essential aspect of collaborative work; in particular, it is a requirement when collaborators need to synchronously and efficiently explore a data visualisation while agreeing on the interpretation of features (i.e., such as during an EDA task) Schneider and Pea [71]. Therefore such exploratory visual analysis requires participants to be aware of each other's visual focus. Previous CSCW studies show that higher mutual awareness of visual focus correlates to higher measures of the perceived quality of collaboration as well as to higher measures of visual coordination D'Angelo and Begel [20], Schneider and Pea [71]. One indicator of mutual awareness is visual coordination: this can be measured by the time a pair of participants concurrently look at the same target (i.e., concurrent gaze pointing behaviour) [54]. However, in our study, we did not use eye trackers as most low-cost VR HMD do not feature them. Instead, we use head-tracked behaviour, which has been reported by several studies as a good proxy for eye movements [10, 57, 79].

2.6 Self-Awareness of Visual Attention Cues

Gaze behavioural cues have been explored in physical manipulation tasks in an MR setup by Jing et al. [39]. This study is particularly interesting as it uses a bi-directional approach to the visualisation of gaze, which is not used in studies that implement the FoV frustum visualisation (e.g, [59]). Their findings highlight how their visualisation increases gaze awareness and joint focus (i.e., visual coordination). In addition, they extend previous research by exploring self-awareness of visual cues. Such exploration leads to interesting results: they report that self-awareness in gaze visualisations increased confidence and certainty that their gaze location is being accurately delivered. With our study, we want to understand the impact of self-awareness of FoV visualisation on confidence and task performances.

2.7 Deictic Pointing

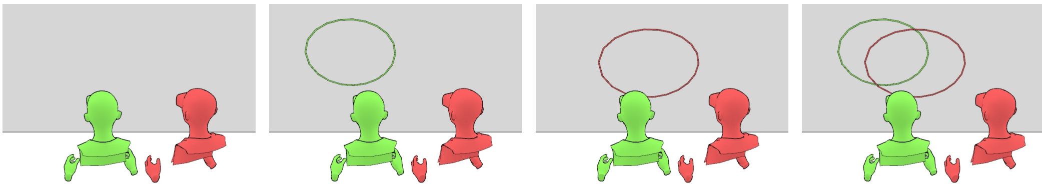
Mutual awareness of visual focus also affects some aspects of communication during collaborative work. For example, when collaborators perform pointing actions during verbal communication (e.g, deictic pointing actions), the awareness of the other's visual focus is an important element. Awareness of the other's visual focus can confirm if the pointing action has been seen by the intended observer and if the target of the pointing action has been understood. Previous work by Wong and Gutwin [81] outlines four stages of deictic pointing action: mutual orientation, preparation, production and holding, and mutual understanding. In these four stages, mutual awareness of visual focus plays an important role both in the *mutual orientation* phase and the *holding and mutual understanding* phase. In the *mutual orientation* phase, the producer of a gesture needs to understand if the intended observer will be able to see the gesture and the target of the pointing gesture. In the *holding and mutual understanding* phase instead, the producer of a gesture needs to hold the gesture until he understands that the observer saw the gesture and understood the correct location of the target. So, mutual awareness of visual focus is a requirement not only for visual coordination during collaboration but also for aspects related to pointing based communication during collaborative work.

2.8 Exploratory Data Analysis

EDA is the process of performing an initial analysis of a dataset, with the aim of identifying interesting patterns and anomalies, elaborating hypotheses and validating assumptions [77]. Such a general and broad task definition has been chosen as it can be generalised to many types of datasets and domains. We specifically choose two types of data-sets: Movies [14, 38] and Cars [70] following the guidelines about the familiarity of the dataset attributes discussed by Saket et al. [67]. To avoid benchmark tasks and to perform a fair evaluation we got inspired by North [53] which suggests using the insights number and quality as performances ratings.

3 STUDY

The purpose of the study is to validate the proposed CoV design, understand what phases of collaborative EDA tasks benefit or are affected by the proposed CoV visualisation, and investigate how the self-awareness of CoV (Figure 3d) affects collaboration. We defined three conditions: a baseline condition in which no CoV is visualised (Figure 3a), a condition representing the already explored use of the CoV in which a user can only see the other participant's CoV but not his own (Figure 3b mono-directional visual cue) and finally a condition in which each participant can see the other participant's CoV as well as his own CoV (Figure 3d bi-directional visual cue).



(a) No CoV (b) One CoV (red view) (c) One CoV (green view) (d) Two CoV

Fig. 3. Experiment one Conditions. (a) No CoV: participants cannot see any CoV. (b)(c) One CoV: each participant can see the CoV of the collaborator but not their own. This visual cue is mono-directional as the collaborator can only see it. (d) Two CoV: participants can see their CoV and the collaborator CoV. This visual cue is bi-directional as it can be seen both by the user and the collaborator.

3.1 Measures

To compare CoV conditions with the baseline we measured visual coordination and mutual awareness of visual attention. Visual coordination was measured by recording the behaviour of participants' heads and calculating the rate of overlap of their field of view. This reflects how long participants spent pointing their heads concurrently toward the same target as used in collaborative studies by Cordeil et al. [19]. To measure how FoV frustum visualisation increases mutual awareness during the collaborative stages, we also measured a self-reported score of mutual awareness and attention allocation [11]. In addition, we used the Nasa TLX questionnaire [32] to measure the load of the task changes across the CoV conditions. We explored how the CoV impacts deictic pointing actions, we measured mutual understanding of spatial references as well as the number of pointing actions performed by each participant. We measured mutual understanding of spatial references by administering an adapted subset of the co-presence questionnaire proposed by Biocca et al. [11]. We examined how the CoV impacts verbal communication by measuring the amount of verbal communication [19] and the balance of verbal communication across participants [23].

To understand how CoV conditions impact the performances of EDA tasks, we recorded the number of the insights and classify them by validity North [53]. Participants were instructed to audio record insights by pressing a button on the controller before verbalising the insight and then pressing the same button to terminate the recording. All audio recordings were transcribed for further analysis. Whenever participants mention more than one fact extracted from the visualisation, we split them unless they were comparing two facts or forming an interpretation. We only counted correct insights for the analysis.

3.2 Apparatus

3.2.1 VR Environment. To keep the application development simple and to avoid noise caused by differences across VR HMDs, we decided to target only Oculus Quest and Quest 2 headsets. The selected VR headsets are six degrees of freedom (DoF) and untethered. We selected these headsets because of their popularity and low retail price. We developed an application for collaborative visual analysis of 3D data using Unity (version 2018.4.14f1) and the Oculus Unity SDK. The application enables displaying visualisations on web pages. Also, it allows each participant to join a real-time session in which other participants' presence is represented by an avatar provided by the Oculus Avatar SDK as shown in Figure 4 (e). Each participant in the VR space is free to move in any direction using the thumbstick controller or physically move using the six DoF of the VR HMD. Avatar movements are streamed via the network (using the Photon Unity Network API), so their behaviour (head and hand movements) and position in the virtual space are reproduced with low latency. The

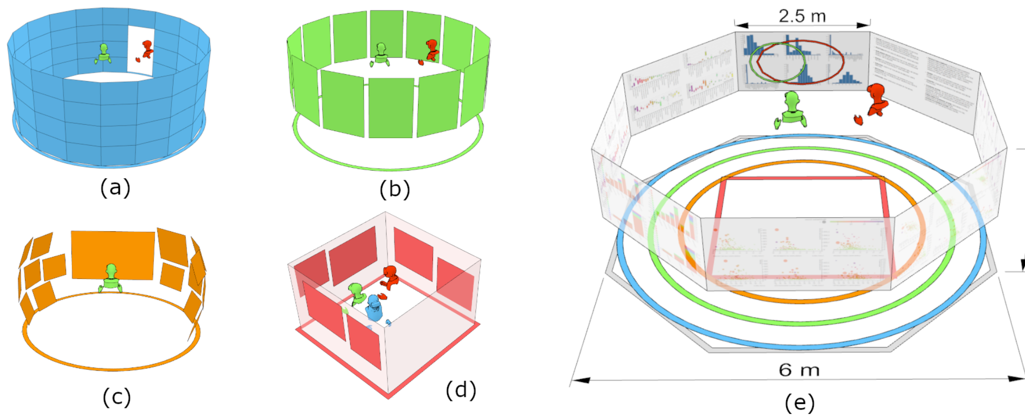
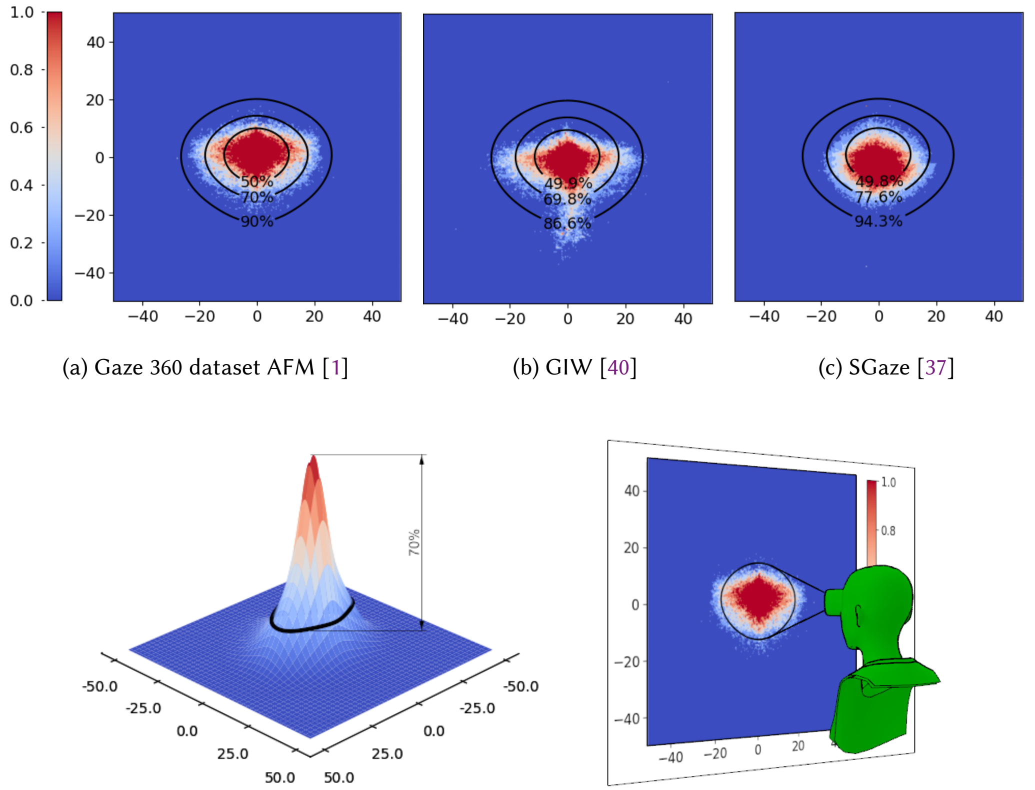


Fig. 4. Inspired by Hybrid Reality Environment for Collaborative Information Analysis used by Cavallo et al. [17] image (b) and Febretti et al. [26] image (a) and in line with cross virtually analytics studies Lee et al. [43] image (d) and Satriadi et al. [69] image (c) we developed an ICVE for the experiment (Sec 2.4). (e) The ICVE is egocentric and consists of a series of 8 screens which are organised in a cylindrical layout around the participants defining the boundaries of the space. The image depicts the ICVE in which two participants can be seen collaborating while trying to complete the exploratory data analysis task. On the base of image (e) is possible to see a size comparison between previously explored layouts and our virtual space.

application also enables participants to talk to each other using the embedded microphone and speakers of the VR HMD. Additionally, the setup supports an observer/moderator to be present (i.e., being heard but not seen) during the VR session. The repository of the VR application can be found on a public github repository [1].

3.2.2 Data Visualisations. We developed three visualisation views using three different datasets: Hollywood movie gender bias (Theme A), the success of Hollywood movies (Theme B), and cars (Theme C). Theme A was inspired by The Bechdel Test [9]- a test that evaluates women presence in Hollywood movies by asking whether a movie features at least two named women characters who talk to each other about anything other than men. The data used for theme A was from [14]. Theme B was about the most successful movies of 2019, displaying attributes related to ratings, financial success and available movie attributes (genre, language, etc..), which was from IMDB [38]. Lastly, Theme C displayed data about the risk rating calculated by insurance companies for cars, correlating with general attributes (engine size, fuel type, etc..), which was from UCI machine learning repository [70].Each theme contained 38 visualisations across the seven screens and one screen contained a glossary for the terms. Each theme also contained four chart types: scatter plots, box and whisker plots, histograms, and stacked bar plots. The scatter plot and box and whiskers plot had one interaction technique that allowed participants to click on items to get textual details (i.e., names of movies or details of the boxes). Each dashboard had a consistent number of visualisations and all visualisations were consistent in size. Captions were added under each visualisation to explain the visualisation and the labels used. The repository of the visualisations and dataset can be found can be found on a public github repository [1], and the visualisations are publicly available on the internet [2].

[1]https://github.com/Collaborative-Immersive-Visual-Toolkit/ConeOfVision
[2]https://graphs-for-collaborative-vr.web.app/



(a) Gaze 360 dataset AFM [1]

(b) GIW [40]

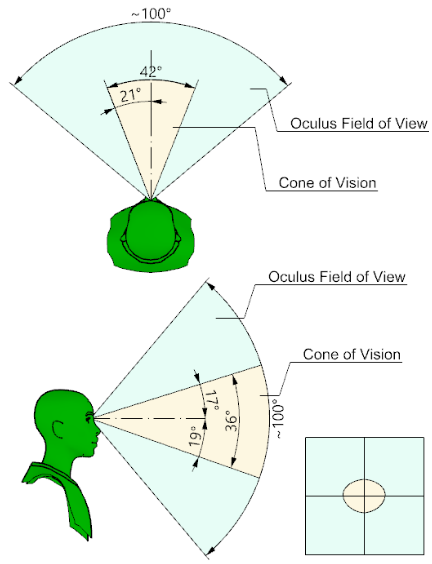
(c) SGaze [37]

(d) Iso-hypses metaphore

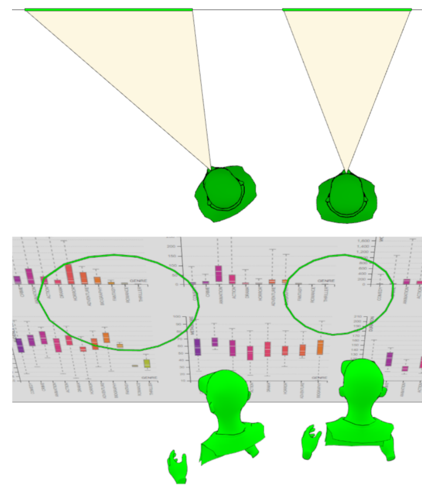
(e) 70% PBC to CoV

Fig. 5. (a) (b) (c) shows the eye distribution in the longitudinal and latitude direction, for each dataset [1, 37, 40] and the amount of eye samples captured by the PBCs generated from [1]. On all three plots, the x-axes show the longitudinal eye angle and on the y-axes the latitudinal eye angle. (d) shows the iso-hypses metaphor where gaze probability is represented as the elevation while the contour is extracted like an iso-hypses. (e) From the identified percentile base contour (PBC) of 70% of the dataset [1], we generate a 3D cone of view.

3.2.3 Cone of Vision. To generate our CoV, we created an average fixation map (AFM) [75] using users' gaze samples from a VR Dataset recorded by Agtzidis et al. [1]. The dataset contains 14 participants watching 360 videos in a VR headset. To generate the average fixation map, we converted all gaze coordinates from a global coordinate system to a head based coordinate system, which represents the latitude and longitude angles that the eye's direction in head coordinates (Figure 5a). The average fixation map is a good approximation of the probability density function (PDF) because of the high number of fixations contained in the dataset. We used PDF approximation to generate percentiles-based contours (PBC). Therefore, the procedure to extract the PBC at the selected percentile uses the PDF gradient to prioritise the areas that have a higher probability of fixations and stop when the count of gaze samples reaches the right percentile. This procedure can be compared to the extraction of iso-hypes from ground elevation information where the iso-hypes are the PBCs, and the elevation information is the gaze probability (as shown in Figure 5d). The contour shape is therefore generated from the data distribution rather than arbitrary chosen



(a) 70% CoV compared to the Oculus Quest FoV



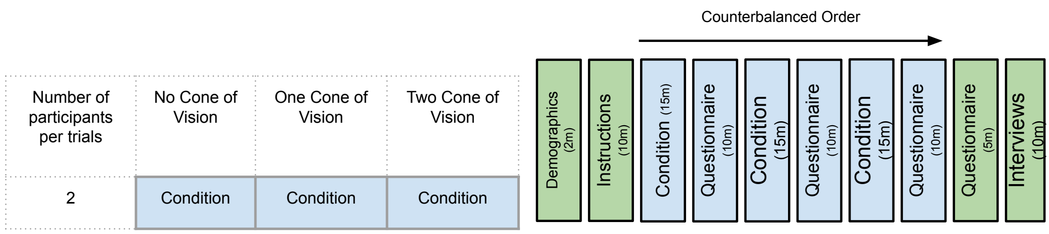
(b) the visualised contour (*conic* curve) changes based on the angle of incidence of the cone with the data

Fig. 6. (a) For comparison, we show the Oculus Quest FoV which consists instead of a ~100° on the longitudinal and a ~100° on the latitudinal. (b) shows how the depicted intersection (green contour) changes according to the angle between the head and the 2D screen.

like in previous FoV visualisations studies [58]. PBC extraction is performed using the library "matplotlib"[3]. After a geometric down-sample, each contour is reduced to a sequence of 20 points with two coordinates. Such percentage-based contours effectively represent the probability of the gaze location in a movement agnostic and context agnostic way. Based on the extracted PBC, we generated our 3D cone Figure 6a. The 3D cone is static in shape but follows the head direction of the user. The 3D cone is not visualised in the ICVE; instead, when the generated cone intersects any of the 2D surfaces present in the scene, we visualise such cone-plane intersection, which is a *conic* curve. We allow only closed conic, and any open conic is closed automatically by following the edges of the 2D screen. Such a conic depends on the angle between the head (cone) direction and plane (as shown in Figure 6b). The intersection is visualised as an elliptically shaped polyline where the colour depends on the user casting the cone. The size of the CoV is $\frac{1}{9}$ of the Oculus FoV (Figure 6a).

As the CoV is generated from datasets where participants were watching panoramic videos in VR [1] there might be important deviations from contours generated from user tests with different tasks and contexts. To understand how well our contours capture visual attention from other scenarios, we evaluate them across two other datasets, one from Hu et al. [37] and the second from Kothari et al. [40]. SGaze dataset contains visual attention data of participants doing a free exploration of a virtual scene with a 6 degree of freedom (DoF) VR HMD [37]. The Gaze in the Wild dataset (GIW) instead contains visual attention data of participants performing everyday tasks while wearing a mobile eye tracker [40]. We extract three PBC (90%, 70%, 50% as shown in Figure 5a) generated from the PDF of the 360 video dataset [1] and we evaluated how much data from GIW and SGaze

[3]https://matplotlib.org/



(a) Experiment Design (b) Experiment procedure

Fig. 7. (a) In the first experiment 12 pairs of participants experience both experimental conditions: Absence of CoV / presence of collaborator CoV / presence of CoV as well as the presence of own CoV (b) Experimental procedure.

they capture. Results show that our PBC captures in SGaze a very similar amount of data (91%, 75%, 51% as shown in Figure 5c), as well for GIW (85%, 69%, 51% as shown in Figure 5b). The repository of the code used to extract the PCB and evaluate them across the dataset can be found on a public github repository [4].

3.2.4 Pilot. While the motion range of the eyes is approximately 50° in any direction for a healthy adult [28], it has been found that they rarely rotate beyond 30° relative to the head [42, 76] and a more comfortable amplitude to observe targets is within 15°[65]. Eyes amplitudes are above 15° during head shifts, instead, lower amplitude matches static target observations [73]. Thus, we searched for the contour that best represents the visual attention that matches the comfortable eye position while observing static targets. Our pilot study involved 3 pairs of participants from our department. We gave participants the option to choose the PBC (90%, 70%, 50%) that best depicted the focus of their visual attention. Participants preference excluded the 90%and 50% PBC because, for the first, they felt it was not focused enough, and for the latter, because they felt that often their visual focus was outside the depicted area. The identified PCB is the one that matched 70% of the data. Consistently with previous works, such PCB is closer to the 15° horizontal eye amplitude, which was outlined as the comfortable eye position. In addition, we tested a 3D representation of the CoV (Figure 2 (c)), assigning transparency of 10% to the lateral surface of the cone. However, such a semi-transparent structure made it harder for the participants to see the data. We did not test the solutions suggested by Hindmarsh et al. [34], Piumsomboon et al. [59] as shown in Figure 2 (a-b), because the size of the FoV was greater than the already excluded 90% PBC shown in Figure 6.

3.3 Participants

Across 90 days we recruited 20 participants (6 women, $M_{\text{Age}} = 30.5$, $SD_{\text{Age}} = 5.9$) online via a number of platforms. We advertised the study on Oculus Quest 2 public Facebook group [5], Oculus subreddit community pages [6], XR Distributed Research Network XRDRN[7], and Prolific[8]. We selected these platforms to target a population interested in VR and could potentially benefit from the CoV feature. We performed a screening with the following inclusion criteria: owning or having access to an

[4]https://graphs-for-collaborative-vr.web.app/
[5]https://www.facebook.com/groups/oculusquest2
[6]https://www.reddit.com/r/oculus/
[7]https://www.xrdrn.org/
[8]https://www.prolific.co/

Oculus Quest 1 or 2 (since we run the study remotely), full proficiency in spoken English, and normal or corrected to normal vision.

Nine participants declared to be heavy VR users, five moderate users, and six to be occasional users. Seventeen participants used Oculus Quest 2 and three participants used Oculus Quest 1. An additional requirement was to be familiar with the chart type we used for the EDA, which consisted of histograms, bar plots, stacked bar plots, and scatter plots. We measured chart familiarity with a preliminary screening questionnaire that contained binary responses (yes or no). All participants had at least a bachelor's degree. Participants received a voucher equivalent of 20$ for a 90 min study. As an incentive, we encouraged participants to perform their best to receive a bonus voucher equivalent of 30$ given to the pair of participants that reported the highest number of valid insights. Moreover, we encouraged participants to collaborate by mentioning that the chance of winning the bonus is shared between the pair, and therefore, each insight reported should be double-checked with the other participant.

3.4 Study Design

The study follows a within-group design, with three conditions (Figure 7). Participants were embodied in an avatar and used a hand pointer to reference the observed dataset. We counterbalanced the order of the conditions to avoid participant learning effects biasing the study results. Also, we counterbalanced the order of the visualisation themes (Section 3.2.2) used in each condition to avoid any difference in data complexity biasing the results.

3.5 Procedure

Participants had to fill out a screening questionnaire before joining the study. We contacted participants via email to gain their consent and arrange a time for the study. We carried the experiment remotely using both web conference and the VR application developed. We started the study with a web conference to give an overview of the study procedure and the three visualisation themes. Next, we asked participants to perform an EDA task, which is extracting insights from the displayed visualisations. We provided participants with examples of valid insights. Next, we explained how to move in the virtual environment, record insights, and use the hand pointer. Once instructions were clear, participants were asked to wear the Oculus Quest, connect to the virtual environment, and start the collaborative task. Between each VR trial (i.e., experimental condition), we asked participants to complete the mutual awareness questionnaire, task load, and usability questionnaire (Table 1). At the end of the experiment, we conducted semi-structured interviews with each pair of participants. In the interviews, participants were asked to report cases in which the CoV helped with the assigned task and cases where it did not (table 4.2). The total time spent on the experiment lasted between 1 hr and 7.3 min and 1 hr and 32.7 min (M = 1 h 20 m, SD = 12.7) and the trial duration lasted between 10 min and 16 min (M = 13 m, SD = 3m). The stop condition was reached when participants declared to have completed the work (i.e. saturation of insights) or when the time was up (13 minutes) and urged them to complete the last insights (i.e. time limitation). This study has been approved by the UCL Interaction Centre (UCLIC) Research Department's Ethics Chair.

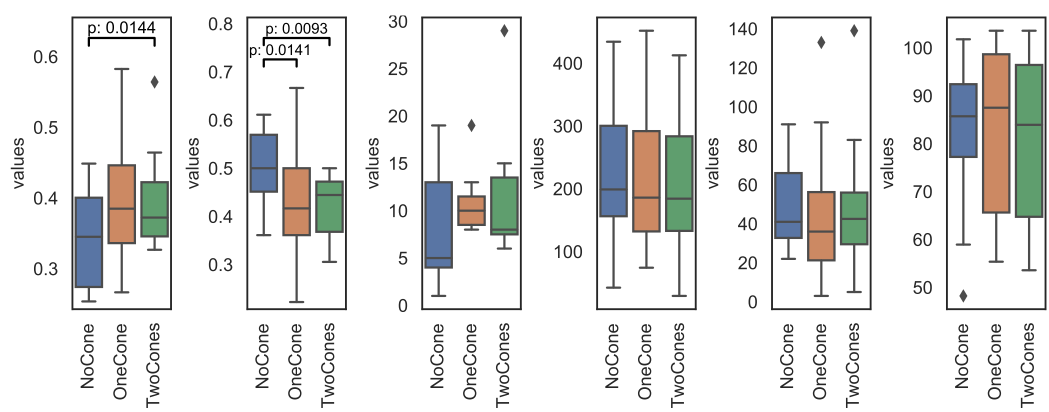
4 RESULTS

4.1 Statistical Analysis

To analyse the collected data, we performed Friedman tests across the three experimental conditions. Significance was set at 0.05, and the Nemenyi test was used for post-hoc analysis where required. In summary, we found significant differences in the average overlap of CoV (Figure 8c), workload

Table 1. Questionnaires

Between conditions questionnaire		
Mutual awareness of visual focus	Q1: I was often aware of the other visual focus in the environment Q2: The other was often aware of my visual focus in the environment	Likert scale agreement 7
Mutual understanding of spatial references	Q3: I understood the exact location of the other's spatial references Q4: The other often understood the exact location of my spatial references	Likert scale agreement 7
Embodiment	Q5: I watched the other participant's avatar to understand what they were doing Q6: The other person watched my avatar to understand what I was doing Q7: I spoke Face to face with the other's avatar	Likert scale frequency 7
Nasa TLX	Q7-Q13 [32]	
System Usability Scale	Q13-Q23 [22]	



(a) Vis. Coupling (b) Nasa TLX (c) Insights (d) Speaking time (e) Pointer Usage (f) SUS

Fig. 8. Behavioural/Questionnaires Results

(Nasa TLX) (Figure 8b), self-reported mutual awareness of visual focus (Figure 9b, Figure 9a), self-reported mutual understanding of spatial references (Figure 9c, Figure 9d), and embodiment usage (Figure 9f, Figure 9e).

4.1.1 Visual Coupling. Visual coupling (Figure 8c and Table 2) is a measure of visual coordination, giving us an idea of how visually synchronised the two participants were throughout the experimental condition. The measure is expressed in a unit scale where 1 means that their head direction was perfectly synchronised and 0 means they never pointed their head towards the same target. As shown in Figure 8c results show a significant difference in visual coupling between No CoV ($M_{\text{NoCoV}} = 0.34$, $SD_{\text{NoCoV}} = 0.07$) and Two CoV ($M_{\text{TwoCoV}} = 0.4$, $SD_{\text{TwoCoV}} = 0.08$). Therefore, the presence of the two CoVs lead to a significant 4% increment in visual synchronisation in the field of view overlap compared to no CoV.

Table 2. Statistical Analysis

	Friedman test		Nemenyi test	
	Chi-Square	p-value	condition	p-value
(a) Visual coupling	14.4645	0.0285*	No CoV, One CoV	0.0781
			No CoV, Two CoV	0.0144*
			One CoV, Two CoV	0.9974
(b) Nasa TlX overall score	9.0	0.0111*	No CoV, One CoV	0.0141*
			No CoV, Two CoV	0.0093**
			One CoV, Two CoV	1.0
(c) Number of insight reported	4.16	0.124	-	-
(d) Speaking time	6.9905	0.033	-	-
(e) Pointer Usage	26.6	0.774	-	-
(f) System Usability Scale Overall score	15.0052	0.0257	-	-

* p < .05, ** p < .01, *** p < .001

4.1.2 NASA TLX Overall Score. We analyse the NASA TLX overall score (unweighted version [32]) by performing a Friedman test $\chi^2 = 9.0$, $p = .0111$. Since the resulting p-value is significant, we compare each condition with a post-hoc test (*Nemenyi*). The test shows a significant difference between No CoV ($M_{\text{NoCoV}} = 0.5$, $SD_{\text{NoCoV}} = 0.07$) and One CoV ($M_{\text{OneCoV}} = 0.42$, $SD_{\text{OneCoV}} = 0.07$) and between No CoV and Two CoVs ($M_{\text{TwoCoV}} = 0.42$, $SD_{\text{TwoCoV}} = 0.07$), but no significant difference between One CoV and Two CoVs. Therefore, the overall Nasa TLX value of the No CoV is statistically higher than the One CoV condition and the Two CoVs condition. The presence of the CoV leads to a 5% reduction in the reported task load (see Figure 8b and Table 2).

4.1.3 Number of Insights Reported. We collected 164 insights reported by pairs of participants and statistically compared them by condition. The analysis of the results did not highlight any significant difference across the experimental conditions as shown in (Figure 8d, Table 2).

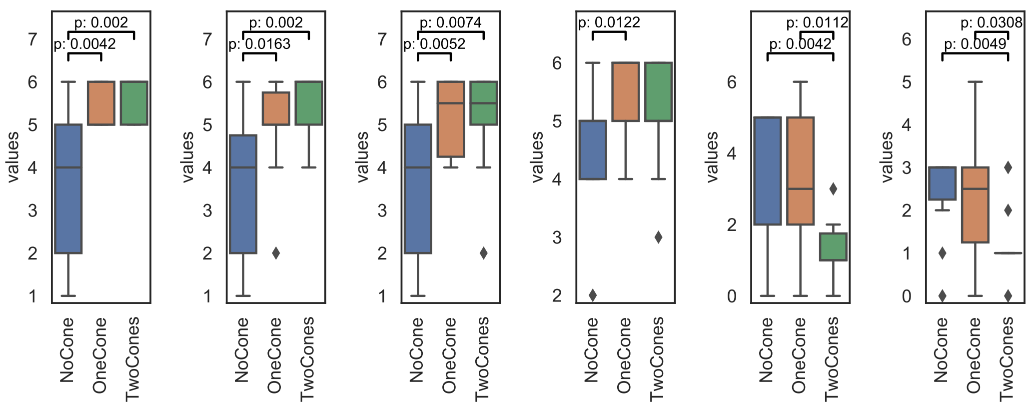
4.1.4 Speaking Time. Verbal communication is an essential part of collaborative dynamics [19]. Using the HMD's microphones, We measured the amount of time participants spent speaking, to see if the experimental conditions affect in any way how much participants communicate. Our results show no significant differences across conditions (Figure 8a and Table 2).

4.1.5 Pointer Usage. We measured the number of pointing actions that participants performed during the trial. The analysis did not highlight any significant difference across the experimental conditions (Figure 8e and Table 2).

4.1.6 System Usability Scale. We analysed the System Usability Scale overall score [44]. The analysis did not highlight any significant difference across the experimental conditions (Figure 8f and Table 2).

4.1.7 Mutual Awareness of Visual Focus (Q1 and Q2). As part of the questionnaires, we asked participants to rate with a 7 point Likert scale (0 to 6) the following statements *Q1 I was often aware of the other visual focus in the environment* (Table 1 and Figure 9a), and *Q2 The other was often aware of my visual focus in the room* (Table 1 and Figure 9b). We analysed the variance with a Friedman test, and since the resulting p-value is significant, we compared each condition with a Nemenyi test (Table 3). The presence of the CoV leads to an increment of the perceived awareness of visual focus from undecided/slightly agree to agree/agree strongly. In particular, we can say that the increment is stronger for the two CoVs (i.e., self-aware CoV) conditions for both Q1 and Q2.

4.1.8 Mutual Understanding of Spatial References (Q3 and Q4). As part of the questionnaires, we asked participants to rate the following statements *Q3 The other often understood the exact location of my spatial references* (Table 1 and Figure 9d) and *Q4 I understood the exact location of the other's spatial references* (Table 1 and Figure 9c). We analysed the variance with a Friedman test Q3 and Q4, since the resulting p-value is significant, we compared each condition with a Nemenyi test (Table 3). The perception of being understood shows an increment only for the One CoV condition while the perception of understanding the other increment equally whenever the CoV is used.



(a) Q1 (b) Q2 (c) Q3 (d) Q4 (e) Q5 (f) Q6

Fig. 9. Questionnaires answers

4.1.9 Embodiment (Q5 and Q6). As part of the questionnaires, we asked participants to rate the following statements *Q5 The other person watched my avatar to understand what I was doing* (Table 1 and Figure 9d) and *Q6 I watched the other participant's avatar to understand what they were doing* (Table 1 and Figure 9c). We analysed the variance with a Friedman test for Q5 and Q6, and since the resulting p-value is significant for both, we compare each condition with a Nemenyi test (Table 3). When Two CoVs are present (i.e., self-aware CoV), we registered a decrement in how much the participant self-report observing the other avatar to understand what the collaborator is doing. Simultaneously we registered a decrement in how much participants perceived the collaborator observing their avatar.

4.2 Qualitative Analysis

The post-experiment semi-structured interviews were audio-recorded and fully transcribed. We coded all transcriptions; codes were initially drawn from research questions and then supplemented with those from the interviews. We group the 42 codes into 4 themes; we report these themes in each of the following subsections.

4.2.1 The Cone of Vision Increases Mutual Awareness . The ability to see each other's CoV increases participants' mutual awareness of visual attention compared to the condition in which the CoV is not present. This means that they can follow each other better when moving their attention from one visualisation to another; they were also able to better switch between collaborative and individual work. *Helped pacing collaborative analysis:* Participants found that seeing the other's

Table 3. Statistical Analysis of Questionnaires

	Friedman test		Nemenyi test	
	Chi-Square	p-value	condition	p-value
Q1: I was often aware of the other visual focus in the environment	18.0476	0.0001***	No CoV, One CoV	0.0042**
			No CoV, Two CoV	0.0020**
			One CoV, Two CoV	0.1797
Q2: The other was often aware of my visual focus in the room	15.6	0.0004***	No CoV, One CoV	0.0163**
			No CoV, Two CoV	0.002**
			One CoV, Two CoV	0.1605
Q3: I understood the exact location of the other's spatial references	12.4615	0.0019**	No CoV, One CoV	0.0052**
			No CoV, Two CoV	0.0074**
			One CoV, Two CoV	0.9141
Q4: The other often understood the exact location of my spatial references	6.8205	0.033*	No CoV, One CoV	0.0122*
			No CoV, Two CoV	0.1006
			One CoV, Two CoV	0.3804
Q5: I watched the other partecipant's avatar to understand what they were doing	12.1333	0.0023**	No CoV, One CoV	0.3822
			No CoV, Two CoV	0.0042**
			One CoV, Two CoV	0.0112*
Q6: The other person watched my avatar to understand what I was doing	10.1052	0.0257*	No CoV, One CoV	0.3822
			No CoV, Two CoV	0.0049**
			One CoV, Two CoV	0.0308*

* $p < .05$, ** $p < .01$, *** $p < .001$

CoV was beneficial to pace the progression of their collaborative analysis. Since participants can independently choose to focus on the same area or focus on different areas, they may split up without both being fully aware of that. For example, one of the two participants could move to another area without notifying the other participant, which leads to one of them being behind without being aware of the split. On the contrary, one might move to another area expecting the other to follow him. Participants reported that when the CoV was present they would not unknowingly split: *"it was useful to see where the other would progress... because I could see sort of when P9 vision started to move on to the next series of graphs [P10]"*, *"If we were doing the same thing and he'd moved on I don't know if I would have caught it straight away without CoV [P15]"*.

The CoV plays an important role in enabling participants to keep their attention coupled.In other words, CoV's presence helped avoid splitting attention because of the lack of communication. The CoV allows participants to be aware of the other person's head movement. Therefore, participants are able to know if the other person was following their progression to the next visualisation. In fact, participants overwhelmingly reported how the benefit of the CoV consisted of being able to follow each other during collaboration: *"I was following his CoV more than I was following his pointer [P8]"*, *"I found the CoV useful in terms of continuing to make sure that we were focusing on the same visualisation [P9]"*.

Beyond showing vision spatial attention, CoV notifies when visual attention is shifting: *" it was useful for two things: be on the same page … but also to pace us in moving so not just the position but also the movements of the cone vision was a useful signal [P12]"*, *"… if I started talking about something I could tell her CoV moved towards mine [P14]"*. Such comments highlight how the CoV is different from the pointer because it is implicit therefore showing behavioural clues which would not normally be deliberate. In particular, some participants reported that such progressions to the next charts were happening without verbally discussing them: *"On some occasions at least I felt* *compelled to look at the same group of graphs that P2 was looking at even if we weren't already talking about them or we did not already discuss to look at those [P3]".*

Participants further reported that such synchronous pacing and progression of their visual attention across the visualised data happened without having to rely on the user embodiment: *"I didn't understand if I didn't look at his avatar unless I had the cone [P10]"*, *"seeing his CoV moving without having to turn and interpret the direction of their avatar made it easier to keep the same pace [P6]".* While to observe the other's CoV depiction, a user keeps facing the data, to observe the collaborator's embodiment, a user needs to look away from the data. Therefore, the CoV allows one to be aware of the other without turning away from the observed data.

Facilitate switching from independent to collaborative work : Participants found that seeing the other's CoV was helpful when switching from independent to collaborative work. During the collaborative session, participants sometimes divided their attention to focus on different goals and then returned to a shared visual context and collaborated towards a common goal. When switching back to collaborative work, the CoV allowed the participant to understand where the other's visual attention was without interpreting the other's embodiment (i.e., interpreting the other head direction). *"when I was looking at something where P1 wasn't looking then I would know that the CoV would help him to actually get on my field of view" [P2]*, *"I think it was useful when I was looking at the glossary and P8 was looking at a visualisation then I could jump exactly to that visualisation that he was on" [P9]*, "like I said to be able to turn around and go like oh that's the direction that he's looking at" [P13], *"I found the cones was helpful to realign our focus" [P6].*

The visual cone depiction, therefore, removed the necessity of looking at the other person's avatar as well as removed any ambiguity about the collaborator's head direction: *"instead of looking at him and tried to triangulate you know I was looking only the cone"[P9]*, *"I didn't have to like look at his avatar and then figure out I could just turn and then I'd see it already and my attention would then gravitate towards that" [P7].*

4.2.2 CoV specificity and hand pointer. Participants found that sometimes the CoV was not specific enough: the CoV provided the general direction of the visual attention but not its exact location. In other words, they highlighted how the CoV shows the general area: *"the cone told me what graph" [P1]*, *"was useful to highlight the graph I was on" [P6]*, *"I knew which window and graph the other was" [P8].* While the pointer would highlight a specific feature: *"... while the pointer told me which column of the plot" [P1]*, *".. but then again, I would use the pointer to increase the focus" [P6]*, *"...buy to point out the particular result I wanted to talk about I had to be more specific" [P8].*

While elaborating on the problem of specificity, participants highlighted how the eyes move independently from the head, therefore, underlying how the CoV may only partially represent the location of the gaze, *"I move my eyes without moving my head so sometimes there was quite a disconnect between the CoV and my gaze"[P7]*, *"it obviously was following tracking of my head rather than my eye " [P3]*; some participants even mentioned eye-tracking as a feature to increase the specificity of the visual clue: *"if you had eye-tracking and you could pinpoint clearer where exactly in the graph I was looking that would be a bit better" [P8].*

Participants also mentioned how the size of the cone (and therefore its specificity) changes depending on their distance from the observed data. When far away from the data visualisations, the cone does feel not informative *"... the cone wasn't always representative of the graph that I was looking at because I was so far back and the cone was so big [P8]".* However, when moving closer the size of the CoV felt more specific/accurate: *"to show that I was looking at a particular boxing whisker or a particular histogram bar I moved forward to narrow the radius [P10]".* Sometimes participants used locomotion to just to make the cone smaller by moving closer to the data. However, when they felt the CoV was not precise enough, most participants used the hand pointer, e.g.: *"I would* *use the pointer rather because that was more focused" [P1]*, or *"I would use the pointer to increase the focus" [P6]*. Participants also highlighted that the familiarity with the hand pointer made it easy to use: *"everybody knows the pointer, so it was in general fantastic [P5]"*, *"when like in meetings I'm often using a laser pointer [P8]"*.

While participants reported using the hand pointer when the cone was not specific enough, they also reported that they used the hand pointer less than in the case in which CoV was not visible. *"We used the pointer more in the first one because we didn't have the CoV [P12]"," we still used it to point out the details absolutely but we used it much much less because we didn't need to [P6]","we didn't have to use the laser pointer every two seconds [P4]"*, *"we would not use the hand pointer a lot [P15]"*, *"I used the pointer less or at least like the impression [P7]"*.

In addition, they reported using the pointer more when their own cone was not visible to them: *"I used the pointer more for that last one [no CoV condition] to give more of an indication of where I'm looking at even though P4 could see where I'm looking at [P8]". "When I could not see my cone, I wasn't sure what the other could see so I used the pointer more to make sure he would understand me [P7]"*

4.2.3 The self-awareness of CoV offers a consistent shared representation. In the self-awareness of the CoV condition, participants reported that they had a better perception of what the other collaborator would see; this allowed them to rely more on the CoV as a behavioural cue. On the contrary, in the other conditions they were unable to clearly depict what the other participant could see: for example, two participants highlighted that because they could not see their own CoV they felt they used the pointer more to make sure the other participant understood what they were talking about: *".. I knew the other could see my CoV, however because I could not see it I would use the pointer more.. [P13]", " ..because we could not see our own cone, we could not rely upon it that much [P15]"*. Another example consisted of two participants being confused by the experimental condition in which their own CoV was not visible. Despite having explained the different experimental conditions to participants in the instructions, two participants believed that because they could not see their own CoV neither their collaborator could see it: *"in the first condition I thought only my collaborators CoV was visible, it took me a while to realise the other could see mine [P2] ", "..yes at the beginning I did not realise my cone was visible, for me it's easier to have a shared reality, so that if he sees my cone I can see it as well so that both cones of vision can be seen [P7]"*. Essentially a consistent representation of the shared reality leaves no doubts.

4.2.4 Confusion and distraction caused by the CoV. If both CoVs are visible, participants can be confused about the ownership of the CoV. To clear their own doubt, participants moved their heads or memorised the colour of their CoV. Participants reported: *"I was a little bit confused which one was mine and which one was hers and I had to move my head [P10]"*, *"I was relying on the colour of the CoV I didn't have any problem distinguish the two of them [P8]"*, *"got a bit confused because I forgot which color I was [P5]"*, *"mine always moved with me so it was difficult to get confused [P9]"*.

One participant found that the collaborator's CoV may distract the collaborator when they are explicitly engaged with the same dataset: *"In that case, I already knew where we were. When we were already looking at the same data and we're talking explicitly about it, so it was for me sometime distracting [P8]"*. Participant also reported that CoV overlapping occurred even if this was negligible: *"the only time that it would overlap is it just when I wasn't paying attention or something like that but it never was the case [P3]"*, *"I mean when we had both our CoV things showing up even in the surveys when I mentioned if it was overlap certain data points it never happened really [P6]"*, *"I don't think it ever got in the way for me... maybe there was a couple of points where the outline of it was overlap a part of the graph I wanted to look at in which case I just moved my head slightly and so it was it wasn't like practically an issue at all [P12]"*, *"The cone never got in the way of what I was doing [P10]"*.

5 DISCUSSION

When the CoV was present, we measured an increment of the awareness of the collaborator's visual attention. Such an increment is more pronounced for the two CoV conditions (Figure 9a). The number of reported insights stayed equal across conditions (Figure 8d) however in the conditions in which the CoV was present the task required a lower cognitive load (Figure 8b, section 4.1.2).

The qualitative analysis offers some explanation for the reduction in cognitive load: we see that the increased awareness of visual focus is most useful when participants switch from individual work to collaborative work (section 4.2.1) and when pacing progression across visualisations (section 4.2.1). CoV visualisations facilitate these two tasks by removing the need of checking and interpret the collaborator avatar's head's direction. The effort saved by users explains the decrease in the overall cognitive task score.

However, the qualitative analysis also highlights how this increment in awareness of visual attention relates only to the general direction of visual attention rather than its focus (section 4.2.2). To overcome such a lack of focus, participants either used the hand pointer or, in very few cases, moved closer to the displayed data to narrow the CoV (section 4.2.2). In fact, we see that the CoV did not influence the laser pointing behaviour during collaboration (section 4.2.2), even if participants perceived to use the pointer less (section 4.2.2).

5.1 Cone of Vision and Collaboration Phases

Piumsomboon et al. [59] suggest that an intelligent system could autonomously choose when to show or hide FoV frustum visualisations based on the detection of collaborative context. However, the previous literature does not explore which phases of collaboration the FoV visualisation should display while suggesting such an idea. Our study identifies three principal collaboration phases during EDA that appear to benefit from CoV visualisations. The first phase consists of switching from independent to collaborative work. A context-aware system could, therefore, temporarily display the CoVs when it detects one participant joining the other for collaborative work. The second phase consists of pacing progression across the visualisations and essentially means the negotiation phase in which participants choose to move to the subsequent visualisation on the displayed dataset. In this case, again, an intelligent system could temporarily display the CoVs when it detects one participant moving to the next visualisation/ next screen. The third phase consists of displaying the CoV when a laser pointing action is performed by one of the participants. In this case, again, an intelligent system could temporarily display the CoVs when a laser pointing action is performed. In this case, the producer of the pointing gesture will be able to understand if the laser pointer is seen or not by the collaborator.

5.2 Cone of Vision Self-Awareness (bi-directional visual attention cue)

Our results also highlight certain benefits concerning the CoV's self-awareness (i.e., Two CoV conditions, bi-directional method Figure 3d). The statistical analysis of the questionnaire responses shows that when participants could see their CoV they reported watching the other avatar less than in the other two conditions (Figure 9e). These results show that CoV self-awareness makes participants feel their collaborator understands what they are doing without watching their avatar. Statistical analysis of questionnaires also shows that participants reported being watched less by their collaborator (Figure 9f) in the self-awareness of CoV condition (i.e., two CoVs conditions). These second results show that they watch the other avatar less in the two CoVs conditions than in the other experimental conditions. Qualitative analysis gives us an insight into why self-awareness of CoV impacts how participants perceived the other one, even if his/her own CoV movements are untangled from the other. One possible explanation is that the CoV perceived reliability is correlated with CoV self-awareness (section 4.2.3). Qualitative analysis confirms our hypothesis that is an extension of Jing et al. [39] from gaze visualisations to FoV visualisations (Section 2.6). Self-awareness of CoV ended up creating a more consistent representation of the shared reality in which both participants see laser pointers and CoVs. Such a representation led to relying more on the CoV visual feedback in general. CoV's self-awareness reduces the number of times participants needed to check each other embodiment. Our results highlight a second benefit concerning the Two CoVs (i.e., CoV's self-awareness), for which we measured a statistically significant 4% increase in visual coupling (Figure 8c, section 4.1.1). Participants spent more time pointing their heads toward the same data when they could see their own CoVs (i.e., approximately 15s on a 6 min task time). Both the qualitative analysis and questionnaires analysis give us ideas about how the CoV's self-awareness caused an increment in visual coupling. The questionnaire's statistical analysis shows us that in the two cones of vision, the need of checking the collaborator's avatar to understand what is doing is lower (Figure 9e). The time saved by avoiding checking and interpreting the collaborator avatar during EDA progression could be the reason for the increment in head direction overlap. By comparing mono-directional (i.e. One CoV) and bi-directional (i.e. Two CoVs) of visual attention cues we are able to outline the benefits of the second compared with the first and therefore we extend the work of Jing et al. [39].

5.3 Deictic Pointing

An important implication for design is that when integrating the CoV in applications, it should not be considered a replacement for pointing. Our results highlight that CoV and laser pointing complement each other, especially during deictic pointing. Indeed, when CoV was present, participants reported that their understanding of spatial references, either verbal or pointing actions, improved with the use of CoV in particular in the conditions of self-awareness of CoV (Figure 9d, section 4.1.9). We can explain such increment because the CoV is a continuous spatial reference (i.e., continuously depicting the general area of visual attention). As such, the CoV also supports a specific phase of hand pointing gestures. Indeed, Wong and Gutwin [81] described how awareness of visual focus plays a role in the different phases of deictic pointing actions (see also 2.7). The CoV supports deictic pointing actions by enabling the producer of the gesture to evaluate if the observer will be able to see it (i.e., collaborators CoV over the gesture's intended target). Additionally, the CoV can support the holding phase of deictic pointing: consisting in holding a pointing gesture until the intended observer has seen it (i.e., laser pointer within observe's CoV). The benefit, in this case, is for the producer of a gesture which is able to understand if a gesture will be seen or has been seen without looking at the observer's avatar or hearing a verbal confirmation from the observer.

5.4 Remote VR studies difficulties

COVID-19 pandemic imposed many VR researchers to shape their experiment workflows to be run remotely. These changes come with several difficulties; for example, researchers need to leverage consumer equipment [74]. On the other hand, there are opportunities, such as developing new toolkits for remote VR experiments [13, 27, 80] and social VR platforms methodologies for remotely running VR studies [63, 66]. However, in our study, we could not use VR social platforms due to the limitation of data serialisations [66]. Our choice did not take advantage of the easy enrolment achieved in VR social platforms, and we registered a high number of no-shows. Non-attendance (no-shows), is a real problem in-lab studies, and remote studies are known to be affected by even higher rates than in-person experiments [72]. Moreover, by using popular low-cost consumer VR HMDs, we managed to run the remote experiment at the cost of a limited data type collection as the eye-tracking capabilities were missing.

6 LIMITATIONS AND FUTURE WORK

While currently, eye trackers can fail due to slippage, spectacles, calibration errors, eye traits of different ethnic groups, this problems may eventually be solved. In the same way, in the years to come, eye sensors might quickly become standard in commercial headsets, which could negate the innovation of CoV as a first choice visualisation method. This could be a limitation for our method, even if we can not be sure that all VR HMDs in the future will include an eye-tracker. A future follow-up study could explore the asymmetric collaborative visual attention cues between eye-tracker-less HMDs (CoV/FoV visualisations) and eye-tracking-ready HMDs (gaze visualisations). Moreover future work could compare compare performances of head based visual cues (CoV/FoV) and Gaze based visual cues in symmetric tasks.

Since XVA's purpose is to exploit visual analytics for seamless integration between different interfaces, another possible follow-up could involve multiple asymmetric scenarios such as HMDs with or without mixed reality capabilities and desktop settings. In this context, we could also explore a comparison of performances with desktop settings (teleconference settings with gaze-awareness).

Another possible limitation of our work relates to the observed data. We designed our experiment to be carried out in an environment where 2D data is displayed on 2D surfaces oriented in 3D space and ordered to form a egocentric and cylindrical room layout [17, 26, 43, 69]. A radically different configuration or a 3D visualisation (e.g., 3D reconstructed environment or a 3D scatter-plot, etc..) might return entirely different results. In these cases of complex 3D data, the intersection of the CoV with the data could not be visible to collaborators because of data obfuscation or might be fragmented and create visual clutter (e.g., as it is not a neat intersection between a cone and a plane). Moreover, such clutter could be greater when one sees his CoV.

A further limitation is related to the number of users and therefore the number of visualised CoV. Previous work by Nguyen and Duval [50] voices a concern related to the use of FoV frustum by many collaborators, which working in narrow spaces may confuse users. In our study, we highlight minor problems related to the CoV generating confusion or the other person's CoV overlapping/obscuring data (Sec 4.2.4). These minor problems could potentially get amplified as the number of users increases. Further work could explore the use of CoV within a large group of users, asynchronous tasks and the relationship between the layout and the CoV to understand how generalisable it can be. Further work could also explore methods to mitigate CoV data overlap when the number of users grows. Ultimately, future work could explore different methods to improve the specificity of the CoV for the usage in low-cost headsets with no eye tracking.

7 CONCLUSION

In this paper, we report on a within-group remote VR study exploring the role of FoV visualisation in ICVEs. We developed CoV, a version of FoV visualisation tailored for 2D VR screens and directly generated by visual attention data. We tested it in a VR virtual office ICVE with 20 participants in the context of collaborative data exploration tasks over 2D visualisations of three datasets.

Our findings indicate that gaze awareness and frustum visualisation foster VR collaboration during EDA tasks by increasing mutual awareness of visual focus. The general benefit of the CoV consists of reducing the number of times when participants need to check each other embodiment and evaluate the collaborator's general direction of attention. Therefore, participants can keep their attention on the visualised data. Our results also highlight how the proposed CoV supports two collaborative phases in VR offices: such as shifts from independent to collaborative work, and aid pacing of progression of EDA across the visualised data. These results extend those of prior work [8, 59, 61] to a new and increasingly popular scenario (i.e. 2D data in VR offices).

Moreover, we propose a novel condition that consists of the "self-CoV" visualisation. Such condition results show an increment in visual coordination and, therefore, it performs better in our scenario. Self-awareness of CoV increases confidence in the visual attention cue amplifying the positive effects described. Our study extends the work of Jing et al. [39] by identifying the same findings from gaze visualisation to FOV visualisations and by comparing mono-directional visual cues to bi-directional visual cues.

Furthermore, our findings show that CoV complements laser pointing by exposing the gesture pointer inside or outside the CoV beforehand and until the gestures' completion, even without verbal confirmation. Such exposure enables the gesture producer to evaluate if the observer can see the performed gesture. This outcome extends previous work about deictic pointing in VR [81, 82].

Given the increasing popularity of ICVEs that use 2D content displayed on VR screens, we hope this work will motivate the research community to study collaborative visual attention cues specifically tailored for this type of 2D data in VR. We hope that this work will also encourage the research community to investigate head-based visual attention cues for eye-tracker-less VR HMDs or as a fallback method for eye-tracking failure.

REFERENCES

[1] Ioannis Agtzidis, Mikhail Startsev, and Michael Dorr. 2019. *360-degree Video Gaze Behaviour: A ground-truth data set and a classification algorithm for eye movements.* Technical Report. 1007–1015 pages. https://doi.org/10.1145/3343031.3350947

[2] Sunggeun Ahn, Jeongmin Son, Sangyoon Lee, and Geehyuk Lee. 2020. Verge-it: Gaze interaction for a binocular head-worn display using modulated disparity vergence eye movement. *Conference on Human Factors in Computing Systems - Proceedings* (2020), 1–7. https://doi.org/10.1145/3334480.3382908

[3] Thomas Alsop. 2021. Virtual reality (VR) headset unit sales worldwide in 4th quarter 2019 and 4th quarter 2020, by device. https://www.statista.com/statistics/987701/vr-unit-sales-brand/

[4] Anand Agarwala; Jinha Lee. 2021. Spatial.io. https://spatial.io/

[5] Chris Anton and Rasmus Larsen. 2021. meetinvr. (2021). https://www.meetinvr.com/

[6] Rowel Atienza, Ryan Blonna, Maria Isabel Saludares, Joel Casimiro, and Vivencio Fuentes. 2016. Interaction techniques using head gaze for virtual reality. *Proceedings - 2016 IEEE Region 10 Symposium, TENSYMP 2016* (2016), 110–114. https://doi.org/10.1109/TENCONSpring.2016.7519387

[7] Per Bækgaard, John Paulin Hansen, Katsumi Minakata, and I. Scott MacKenzie. 2019. A fitts' law study of pupil dilations in a head-mounted display. *Eye Tracking Research and Applications Symposium (ETRA)* (2019). https://doi.org/10.1145/3314111.3319831

[8] Huidong Bai, Prasanth Sasikumar, Jing Yang, and Mark Billinghurst. 2020. A User Study on Mixed Reality Remote Collaboration with Eye Gaze and Hand Gesture Sharing. *Conference on Human Factors in Computing Systems - Proceedings* (2020). https://doi.org/10.1145/3313831.3376550

[9] Alison Bechdel. 1985. Bechdel Test. https://en.wikipedia.org/wiki/Bechdel_test.

[10] B. Biguer, M. Jeannerod, and C. Prablanc. 1982. The coordination of eye, head, and arm movements during reaching at a single visual target. *Experimental Brain Research* 46, 2 (5 1982), 301–304. https://doi.org/10.1007/BF00237188

[11] Frank Biocca, Chad Harms, and Jenn Gregg. 2001. The networked minds measure of social presence: Pilot test of the factor structure and concurrent validity. *4th Annual International Workshop on Presence* December (2001), 1–9. http://astro.temple.edu/~lombard/ISPR/Proceedings/2001/Biocca2.pdf

[12] Pieter Blignaut and Daniël Wium. 2014. Eye-tracking data quality as affected by ethnicity and experimental design. *Behavior Research Methods* 46, 1 (2014), 67–80. https://doi.org/10.3758/s13428-013-0343-0

[13] Riccardo Bovo, Daniele Giunchi, Anthony Steed, and Thomas Heinis. 2022. MR-RIEW: An MR Toolkit for Designing Remote Immersive Experiment Workflows. In *2022 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*. IEEE, 766–767.

[14] Sharon Brener. 2013. Bechdel test dataset. https://data.world/sharon/bechdel-test/workspace/file?filename=movies.csv.

[15] Wolfgang Buschel, Anke Lehmann, and Raimund Dachselt. 2021. Miria: A mixed reality toolkit for the in-situ visualization and analysis of spatio-temporal interaction data. *Conference on Human Factors in Computing Systems - Proceedings* (2021). https://doi.org/10.1145/3411764.3445651

[16] Simon Butscher, Sebastian Hubenschmid, Jens Müller, Johannes Fuchs, and Harald Reiterer. 2018. Clusters, trends, and outliers: How Immersive technologies can facilitate the collaborative analysis of multidimensional data. *Conference on Human Factors in Computing Systems - Proceedings* 2018-April (2018). https://doi.org/10.1145/3173574.3173664

[17] Marco Cavallo, Mishal Dholakia, Matous Havlena, Kenneth Ocheltree, and Mark Podlaseck. 2019. Dataspace: A reconfigurable hybrid reality environment for collaborative information analysis. *26th IEEE Conference on Virtual Reality and 3D User Interfaces, VR 2019 - Proceedings* (2019), 145–153. https://doi.org/10.1109/VR.2019.8797733

[18] Marco Cavallo, Mishal Dholakia, Matous Havlena, Kenneth Ocheltree, and Mark Podlaseck. 2019. Immersive insights: A hybrid analytics system for collaborative exploratory data analysis. *Proceedings of the ACM Symposium on Virtual Reality Software and Technology, VRST* (2019). https://doi.org/10.1145/3359996.3364242

[19] Maxime Cordeil, Tim Dwyer, Karsten Klein, Bireswar Laha, Kim Marriott, and Bruce H. Thomas. 2017. Immersive Collaborative Analysis of Network Connectivity: CAVE-style or Head-Mounted Display? *IEEE Transactions on Visualization and Computer Graphics* 23, 1 (2017), 441–450. https://doi.org/10.1109/TVCG.2016.2599107

[20] Sarah D'Angelo and Andrew Begel. 2017. Improving communication between pair programmers using shared gaze awareness. *Conference on Human Factors in Computing Systems - Proceedings* 2017-Janua (2017), 6245–6255. https://doi.org/10.1145/3025453.3025573

[21] David Whelan. 2021. engage vr. https://engagevr.io/

[22] Marie Diener-west, Jeff Sauro, and John Brooke. 2020. SUS: A 'Quick and Dirty' Usability Scale. *Usability Evaluation In Industry* (2020), 207–212. https://doi.org/10.1201/9781498710411-35

[23] Joan Morris Dimicco, Anna Pandolfo, and Walter Bender. 2004. Influencing Group Participation with a Shared Display. (2004).

[24] Ciro Donalek, S. G. Djorgovski, Alex Cioc, Anwell Wang, Jerry Zhang, Elizabeth Lawler, Stacy Yeh, Ashish Mahabal, Matthew Graham, Andrew Drake, Scott Davidoff, Jeffrey S. Norris, and Giuseppe Longo. 2015. Immersive and collaborative data visualization using virtual reality platforms. *Proceedings - 2014 IEEE International Conference on Big Data, IEEE Big Data 2014* (2015), 609–614. https://doi.org/10.1109/BigData.2014.7004282

[25] Heiko Drewes, Mohamed Khamis, and Florian Alt. 2018. Smooth pursuit target speeds and trajectories. *ACM International Conference Proceeding Series* (2018), 139–146. https://doi.org/10.1145/3282894.3282913

[26] Alessandro Febretti, Arthur Nishimoto, Terrance Thigpen, Jonas Talandis, Lance Long, J. D. Pirtle, Tom Peterka, Alan Verlo, Maxine Brown, Dana Plepys, Dan Sandin, Luc Renambot, Andrew Johnson, and Jason Leigh. 2013. CAVE2: a hybrid reality environment for immersive simulation and information analysis. *The Engineering Reality of Virtual Reality 2013* 8649, March 2013 (2013), 864903. https://doi.org/10.1117/12.2005484

[27] Martin Feick, Niko Kleer, Anthony Tang, and Antonio Krüger. 2020. The Virtual Reality Questionnaire Toolkit. In *Adjunct Publication of the 33rd Annual ACM Symposium on User Interface Software and Technology* (Virtual Event, USA) *(UIST '20 Adjunct)*. Association for Computing Machinery, 68–69. https://doi.org/10.1145/3379350.3416188

[28] Virgilio F Ferrario, Chiarella Sforza, Graziano Serrao, GianPiero Grassi, and Erio Mossi. 2002. Active range of motion of the head and cervical spine: a three-dimensional investigation in healthy young adults. *Journal of orthopaedic research* 20, 1 (2002), 122–129.

[29] Mike Fraser, Steve Benford, Jon Hindmarsh, and Christian Heath. 1999. Supporting awareness and interaction through collaborative virtual interfaces. In *UIST (User Interface Software and Technology): Proceedings of the ACM Symposium*, Vol. 1. ACM, New York, New York, USA, 27–36. https://doi.org/10.1145/320719.322580

[30] B. Fröhler, C. Anthes, F. Pointecker, J. Friedl, D. Schwajda, A. Riegler, S. Tripathi, C. Holzmann, M. Brunner, H. Jodlbauer, H. C. Jetter, and C. Heinzl. 2022. A Survey on Cross-Virtuality Analytics. *Computer Graphics Forum* 41, 1 (2022), 465–494. https://doi.org/10.1111/cgf.14447

[31] Adam Glazier. 2017. Designing Screen Interfaces for VR (Google I/O '17). https://www.youtube.com/watch?v=ES9jArHRFHQ

[32] Sandra G. Hart and Lowell E. Staveland. 1988. Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. *Advances in Psychology* 52, C (1988), 139–183. https://doi.org/10.1016/S0166-4115(08)62386-9

[33] Keita Higuch, Ryo Yonetani, and Yoichi Sato. 2016. Can eye help you?: Effects of visualizing eye fixations on remote collaboration scenarios for physical tasks. In *Conference on Human Factors in Computing Systems - Proceedings*. Association for Computing Machinery, New York, NY, USA, 5180–5190. https://doi.org/10.1145/2858036.2858438

[34] Jon Hindmarsh, Mike Fraser, Christian Heath, Steve Benford, and Chris Greenhalgh. 1998. Fragmented interaction: Establishing mutual orientation in virtual environments. In *Proceedings of the ACM Conference on Computer Supported Cooperative Work*. ACM Press, New York, New York, USA, 217–226.

[35] Adrian H. Hoppe, Kai Westerkamp, Sebastian Maier, Florian van de Camp, and Rainer Stiefelhagen. 2018. Multi-user collaboration on complex data in virtual and augmented reality. *Communications in Computer and Information Science* 851 (2018), 258–265. https://doi.org/10.1007/978-3-319-92279-9{_}35

[36] Alex Howland. 2021. Virbela. https://www.virbela.com/

[37] Zhiming Hu, Congyi Zhang, Sheng Li, Guoping Wang, and DInesh Manocha. 2019. SGaze: A Data-Driven Eye-Head Coordination Model for Realtime Gaze Prediction. *IEEE Transactions on Visualization and Computer Graphics* 25, 5 (2019), 2002–2010. https://doi.org/10.1109/TVCG.2019.2899187

[38] IMDb. 1990. IMDb Datasets. https://www.imdb.com/interfaces/.

[39] Allison Jing, Kieran William May, Mahnoor Naeem, Gun Lee, and Mark Billinghurst. 2021. EyemR-Vis: Using Bi-Directional Gaze Behavioural Cues to Improve Mixed Reality Remote Collaboration. In *Conference on Human Factors in Computing Systems - Proceedings*. Association for Computing Machinery. https://doi.org/10.1145/3411763.3451844

[40] Rakshit Kothari, Zhizhuo Yang, Christopher Kanan, Reynold Bailey, Jeff B. Pelz, and Gabriel J. Diaz. 2020. Gaze-in-wild: A dataset for studying eye and head coordination in everyday activities. *Scientific Reports* 10, 1 (2020), 1–23. https://doi.org/10.1038/s41598-020-59251-5

[41] Mikko Kytö, Barrett Ens, Piumsomboon Thammathip, Lee Gun A., and Mark Billinghurst. 2018. Pinpointing. *Education + Training* 23, 3 (2018), 88–89. https://doi.org/10.1108/eb016812

[42] Michael F. Land. 2009. Vision, eye movements, and natural behavior. *Visual Neuroscience* 26, 1 (2009), 51–62. https://doi.org/10.1017/S0952523808080899

[43] Benjamin Lee, Xiaoyun Hu, Maxime Cordeil, Arnaud Prouzeau, Bernhard Jenny, and Tim Dwyer. 2021. Shared surfaces and spaces: Collaborative data visualisation in a co-located immersive environment. *IEEE Transactions on Visualization and Computer Graphics* 27, 2 (2021), 1171–1181. https://doi.org/10.1109/TVCG.2020.3030450

[44] James R. Lewis. 2018. The System Usability Scale: Past, Present, and Future. *International Journal of Human-Computer Interaction* 34, 7 (2018), 577–590. https://doi.org/10.1080/10447318.2018.1455307

[45] Yin Li, Alireza Fathi, and James M Rehg. 2013. Learning to Predict Gaze in Egocentric Video. (2013), 3216–3223. https://doi.org/10.1109/ICCV.2013.399

[46] Tahir Mahmood, Erik Butler, Nicholas Davis, Jian Huang, and Aidong Lu. 2018. Building Multiple Coordinated Spaces for Effective Immersive Analytics through Distributed Cognition. *2018 International Symposium on Big Data Visual and Immersive Analytics, BDVA 2018* (2018). https://doi.org/10.1109/BDVA.2018.8533893

[47] Mark McGill, Aidan Kehoe, Euan Freeman, and Stephen Brewster. 2020. Expanding the Bounds of Seated Virtual Workspaces. *ACM Transactions on Computer-Human Interaction* 27, 3 (2020). https://doi.org/10.1145/3380959

[48] Meta. 2020. Infinite Office. https://www.youtube.com/watch?v=ES9jArHRFHQ

[49] David Nguyen and John Canny. 2005. MultiView: Spatially faithful group video conferencing. *CHI 2005: Technology, Safety, Community: Conference Proceedings - Conference on Human Factors in Computing Systems* (2005), 799–808.

[50] Thi Thuong Huyen Nguyen and Thierry Duval. 2015. A survey of communication and awareness in collaborative virtual environments. In *2014 International Workshop on Collaborative Virtual Environments, 3DCVE 2014*. Institute of Electrical and Electronics Engineers Inc., 1–8. https://doi.org/10.1109/3DCVE.2014.7160928

[51] Diederick C. Niehorster, Thiago Santini, Roy S. Hessels, Ignace T.C. Hooge, Enkelejda Kasneci, and Marcus Nyström. 2020. The impact of slippage on the data quality of head-worn eye trackers. *Behavior Research Methods* 52, 3 (2020), 1140–1160. https://doi.org/10.3758/s13428-019-01307-0

[52] Arthur Nishimoto and Andrew Johnson. 2019. Extending virtual reality displaywall environments using augmented reality. *Proceedings - SUI 2019: ACM Conference on Spatial User Interaction* (oct 2019). https://doi.org/10.1145/3357251.3357579

[53] Chris North. 2006. Toward measuring visualization insight. *IEEE Computer Graphics and Applications* 26, 3 (2006), 6–9. https://doi.org/10.1109/MCG.2006.70

[54] Marc-Antoine Nüssli. 2011. Dual Eye-Tracking Methods for the Study of Remote Collaborative Problem Solving. *PhD Thesis, ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE* 5232 (2011). https://doi.org/10.5075/epfl-thesis-5232

[55] Santtu Parikka and Juha Ruistola. 2021. Glue Collaboration. https://glue.work/

[56] Leonardo Pavanatto, Chris North, Doug A. Bowman, Carmen Badea, and Richard Stoakley. 2021. Do we still need physical monitors? An evaluation of the usability of AR virtual monitors for productivity work. *Proceedings - 2021 IEEE Conference on Virtual Reality and 3D User Interfaces, VR 2021* (2021), 759–767. https://doi.org/10.1109/VR50410.2021.00103

[57] Jeff Pelz, Mary Hayhoe, and Russ Loeber. 2001. The coordination of eye, head, and hand movements in a natural task. *Experimental Brain Research* 139, 3 (8 2001), 266–277. https://doi.org/10.1007/s002210100745

[58] Thammathip Piumsomboon, Arindam Dey, Barrett Ens, Gun Lee, and Mark Billinghurst. 2017. CoVAR: Mixed-Platform Remote Collaborative Augmented and Virtual Realities System with Shared Collaboration Cues. In *Adjunct Proceedings of the 2017 IEEE International Symposium on Mixed and Augmented Reality, ISMAR-Adjunct 2017*. Institute of Electrical and Electronics Engineers Inc., 218–219. https://doi.org/10.1109/ISMAR-Adjunct.2017.72

[59] Thammathip Piumsomboon, Arindam Dey, Barrett Ens, Gun Lee, and Mark Billinghurst. 2019. The effects of sharing awareness cues in collaborative mixed reality. *Frontiers Robotics AI* 6, FEB (2019). https://doi.org/10.3389/frobt.2019.00005

[60] Thammathip Piumsomboon, Gun A. Lee, Jonathon D. Hart, Barrett Ens, Robert W. Lindeman, Bruce H. Thomas, and Mark Billinghurst. 2018. Mini-me: An adaptive avatar for Mixed Reality remote collaboration. *Conference on Human Factors in Computing Systems - Proceedings* 2018-April (2018), 1–13. https://doi.org/10.1145/3173574.3173620

[61] Michael Prilla. 2019. "I simply watched where she was looking at": Coordination in short-term synchronous cooperative mixed reality. *Proceedings of the ACM on Human-Computer Interaction* 3, GROUP (2019). https://doi.org/10.1145/3361127

[62] Yuan Yuan Qian and Robert J. Teather. 2017. The eyes don't have it. (2017), 91–98. https://doi.org/10.1145/3131277.3132182

[63] Rivu Radiah, Ville Mäkelä, Sarah Prange, Sarah Delgado Rodriguez, Robin Piening, Yumeng Zhou, Kay Köhle, Ken Pfeuffer, Yomna Abdelrahman, Matthias Hoppe, Albrecht Schmidt, and Florian Alt. 2021. Remote VR Studies: A Framework for Running Virtual Reality Studies Remotely Via Participant-Owned HMDs. *ACM Transactions on Computer-Human Interaction* 28, 6 (2021). https://doi.org/10.1145/3472617 arXiv:2102.11207

[64] Patrick Reipschlager, Tamara Flemisch, and Raimund Dachselt. 2021. Personal augmented reality for information visualization on large interactive displays. *IEEE Transactions on Visualization and Computer Graphics* 27, 2 (2021), 1182–1192. https://doi.org/10.1109/TVCG.2020.3030460 arXiv:2009.03237

[65] Sohrab Saeb, Cornelius Weber, and Jochen Triesch. 2011. Learning the optimal control of coordinated eye and head movements. *PLoS Computational Biology* 7, 11 (2011). https://doi.org/10.1371/journal.pcbi.1002253

[66] David Saffo, Sara Di Bartolomeo, Caglar Yildirim, and Cody Dunne. 2021. Remote and collaborative virtual reality experiments via social vr platforms. *Conference on Human Factors in Computing Systems - Proceedings* (2021). https://doi.org/10.1145/3411764.3445426

[67] Bahador Saket, Alex Endert, and Çağatay Demiralp. 2017. Task-Based Effectiveness of Basic Visualizations. *arXiv* 25, 7 (2017), 2505–2512.

[68] Čeněk Šašinka, Zdeněk Stachoň, Michal Sedlák, Jiří Chmelík, Lukáš Herman, Petr Kubíček, Alžběta Šašinková, Milan Doležal, Hynek Tejkl, Tomáš Urbánek, Hana Svatoňová, Pavel Ugwitz, and Vojtěch Juřík. 2019. Collaborative immersive virtual environments for education in geography. *ISPRS International Journal of Geo-Information* 8, 1 (2019). https://doi.org/10.3390/ijgi8010003

[69] Kadek Ananta Satriadi, Barrett Ens, Maxime Cordeil, Tobias Czauderna, and Bernhard Jenny. 2020. Maps around Me: 3D Multiview Layouts in Immersive Spaces. *Proceedings of the ACM on Human-Computer Interaction* 4, ISS (2020). https://doi.org/10.1145/3427329

[70] Jeffrey Schlimmer. 1985. Automobile data set. http://archive.ics.uci.edu/ml/datasets/Automobile.

[71] Bertrand Schneider and Roy Pea. 2013. Real-time mutual gaze perception enhances collaborative learning and collaboration quality. *International Journal of Computer-Supported Collaborative Learning* 8, 4 (2013), 375–397. https://doi.org/10.1007/s11412-013-9181-4

[72] Scott, S.; Johnson, K. 2020. Pros and cons of remote moderated testing: Considerations for ongoing research during COVID-19. https://boldinsight.com/pros-and-cons-of-remote-testing-considerations-for-ongoing-research-during-covid-19/

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

[74] Anthony Steed, Daniel Archer, Ben Congdon, Sebastian Friston, David Swapp, and Felix J. Thiel. 2021. Some Lessons Learned Running Virtual Reality Experiments Out of the Laboratory. (2021), 5–7. arXiv:2104.05359 http://arxiv.org/abs/2104.05359

[75] Hamed R. Tavakoli, Esa Rahtu, Juho Kannala, and Ali Borji. 2019. Digging deeper into egocentric gaze prediction. In *Proceedings - 2019 IEEE Winter Conference on Applications of Computer Vision, WACV 2019*. Institute of Electrical and Electronics Engineers Inc., 273–282. https://doi.org/10.1109/WACV.2019.00035

[76] Alvin R Tilley et al. 2002. Henry Dreyfuss Associates. *The measure of man and woman: human factors in design* (2002).

[77] Jhon W Tukey. 1993. Análisis de datos exploratorios: pasado, presente y futuro. 0704 (1993), 1–61.

[78] Roel Vertegaal, Ivo Weevers, Changuk Sohn, and Chris Cheung. 2003. GAZE-2: Conveying eye contact in group video conferencing using eye-controlled camera direction. *Conference on Human Factors in Computing Systems - Proceedings* 5 (2003), 521–528.

[79] Peng Wang, Shusheng Zhang, Xiaoliang Bai, Mark Billinghurst, Weiping He, Shuxia Wang, Xiaokun Zhang, Jiaxiang Du, and Yongxing Chen. 2019. Head pointer or eye gaze: Which helps more in MR remote collaboration. In *26th IEEE Conference on Virtual Reality and 3D User Interfaces, VR 2019 - Proceedings*. Institute of Electrical and Electronics Engineers Inc., 1219–1220. https://doi.org/10.1109/VR.2019.8798024

[80] Marcus R Watson, Benjamin Voloh, Christopher Thomas, Asif Hasan, and Thilo Womelsdorf. 2019. USE: An integrative suite for temporally-precise psychophysical experiments in virtual environments for human, nonhuman, and artificially intelligent agents. *Journal of neuroscience methods* 326 (2019), 108374.

[81] Nelson Wong and Carl Gutwin. 2010. Where are you pointing? The accuracy of deictic pointing in CVEs. In *Conference on Human Factors in Computing Systems - Proceedings*, Vol. 2. ACM Press, New York, New York, USA, 1029–1038. https://doi.org/10.1145/1753326.1753480

[82] Nelson Wong and Carl Gutwin. 2014. Support for deictic pointing in CVEs. (2014), 1377–1387. https://doi.org/10.1145/2531602.2531691

[83] Wanze Xie, Yining Liang, Janet Johnson, Andrea Mower, Samuel Burns, Colleen Chelini, Paul D'Alessandro, Nadir Weibel, and Jürgen P. Schulze. 2020. Interactive Multi-User 3D Visual Analytics in Augmented Reality. *IS and T*

International Symposium on Electronic Imaging Science and Technology 2020, 13 (2020), 1–6. https://doi.org/10.2352/ISSN.2470-1173.2020.13.ERVR-363

[84] Difeng Yu and Xueshi Lu. 2021. Gaze-supported 3d object manipulation in virtual reality. *Conference on Human Factors in Computing Systems - Proceedings* (2021). https://doi.org/10.1145/3411764.3445343

Received January 2022; revised April 2022; accepted August 2022