Social Density and its Impact on Behaviour in Virtual Environments

Julie Williamson
julie.williamson@glasgow.ac.uk
University of Glasgow
Glasgow, UK

Silvia Rossi
s.rossi@cwi.nl
Centrum Wiskunde & Informatica (CWI)
Amsterdam, Netherlands

Ross Johnstone
ross.johnstone@glasgow.ac.uk
University of Glasgow
Glasgow, UK

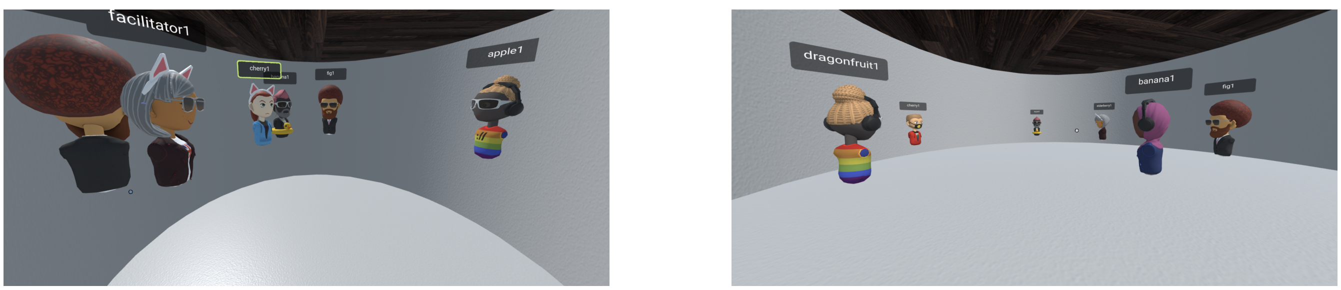
Irene Viola
irene.viola@cwi.nl
Centrum Wiskunde & Informatica (CWI)
Amsterdam, Netherlands

John Williamson
johnh.williamson@glasgow.ac.uk
University of Glasgow
Glasgow, UK

Thomas Röggla
t.roggla@cwi.nl
Centrum Wiskunde & Informatica (CWI)
Amsterdam, Netherlands

David A. Shamma
aymans@acm.org
Centrum Wiskunde & Informatica (CWI)
Amsterdam, Netherlands

Pablo Cesar
garcia@cwi.nl
Centrum Wiskunde & Informatica (CWI) and Delft University of Technology
Amsterdam, Delft, Netherlands



(a) The small room with high social density

(b) The large room with low social density

Figure 1: A large and a small virtual room created spaces with low or high social density.

Abstract

Virtual environments make it possible to connect and collaborate in social immersive realities, but there are still open questions about the influence of their design on the user experience. We conducted a study with 48 participants divided into groups of 6, completing conversational tasks in an instrumented virtual environment. Using a mix-methods approach, combining qualitative and quantitative research methods (interviews, questionnaires, conversation and movement analysis), we compared between two virtual environment designs. We found that the social density (or effective capacity) of the designed virtual environment influenced the quality of interaction between the participants.

CCS Concepts

• **Human-centered computing → Empirical studies in HCI**; **Virtual reality**; **Empirical studies in collaborative and social computing**.

Keywords

Social VR, social density, social signal processing

ACM ISBN 979-8-4007-2051-2/2025/10
https://doi.org/10.1145/3746269.3760421

ACM Reference Format:
Julie Williamson, Silvia Rossi, Ross Johnstone, Irene Viola, John Williamson, Thomas Röggla, David A. Shamma, and Pablo Cesar. 2025. Social Density and its Impact on Behaviour in Virtual Environments. In *Proceedings of the 3rd International Workshop on Interactive eXtended Reality (IXR '25), October 27–28, 2025, Dublin, Ireland*. ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/3746269.3760421

1 Introduction

We are moving towards a future where immersive environments and virtual realities (VR) mediate our social and collaborative interactions. How our interpersonal interactions play out in VR is significantly influenced by how these experiences are designed, but we have limited models for understanding social phenomenon when mediated by immersive environments. *Digital proxemics* is an emerging model for understanding how we use space in virtual environments, focusing on social signals of body position relative to others in virtual space [4, 27, 28]. *Interpersonal social signals* like proximity, mutual gaze, and conversational fluency, which represent the signals created when two or more people interact, provide rich insights into what happens between individuals. Defining these interpersonal social signals and developing more sophisticated models of interpersonal interaction in immersive environments is needed to inform the design of more satisfying and successful communications and collaborations in social VR.

This paper explores the *quality* of interactions with different social densities [21, 26], that is the ratio of the number of people to the size of a space. Social density influences how crowded, cozy, or empty a space feels. Our hypothesis is that manipulating social density will influence how interaction unfolds. Our analysis focuses on the interpersonal social signals that we give off when interacting in immersive environments, focusing on how social density creates observable changes in these interpersonal social signals. We completed a lab-based user study (N=48) where eight groups of six participants completed conversational tasks while wearing head-mounted displays (HMDs) in an instrumented virtual environment. Results show different interaction and conversation patterns, depending on the social density of the virtual environment.

This paper makes the following contributions:

(1) A comparison of two social density conditions (low and high), resulting from a between-subjects evaluation (N=48) of conversational tasks in a virtual environment.
(2) An open source dataset capturing rich social signals including proximity, gesture, gaze, and conversation under low and high *social densities*.

2 Related Work

2.1 Proxemics

Research on proxemics in physical spaces has explored how people use space [12, 16], intended usage of spaces [22], and space design [10, 13]. Hall's foundational work on proxemics [12] provides the baseline for understanding how people use physical space. Factors like body visibility, voice loudness, body heat, odour, and personal preferences all contribute to how people use space, which is dynamic and contextually influenced by culture. Hall categorises four proxemic zones: intimate, personal, social, and public. The *intimate* zone (< 0.46 meters) entails more physical contact, close field of view, and potential awareness of breath or body heat. The *personal* zone (0.46–1.2 meters) allows for physical contact, but not consistently, and facial expressions and visual attention are more noticeable. Transitioning to the *social* zone (1.2–3.6 meters), one would expect less physical contact, if any, and the ability to observe the other person's entire body. Lastly, the *public* zone (> 3.6 meters) makes facial expressions and voice less impactful but enhances peripheral awareness of the immediate environment.

Beyond interpersonal distance, proxemics research also incorporates visual attention and body orientation when people interact in groups. Kendon introduces the concept of *F-formations*, emphasising the relative orientation of a group over time as a significant aspect of proximity [16]. F-formations emerge during social interactions among people in close proximity, who share a space with exclusive, direct, and equal access.

Applying physical proxemics to behaviour in virtual environments raises challenges given the absence of typical social and environmental constraints in the virtual world. Incorporating olfaction and haptics, which play vital roles in physical proxemics, presents difficulties within virtual settings. Additionally, the auditory experience through Head-Mounted Display (HMD) headphones diverges significantly from real-world auditory perceptions [24]. However, Virtual Environments (VEs) offer the intriguing possibility of intentionally manipulating, distorting or enhancing proxemic cues to shape the social dynamics of the virtual space.

2.2 Social Density

Social density refers to the number of individuals present in a given physical [3] or virtual space [21]. Social density is a crucial factor that influences social interactions, personal comfort, and behaviour patterns within space. Whyte's research asks what makes public spaces feel spacious or crowded, exploring the availability and layout of seating, street positioning, and exposure to elements [26]. His work integrates physical space design with human proxemics, examining effective capacity and perceived crowding. Effective capacity, which considers factors like proximity to others, comfort, and amenities, can significantly differ from physical capacity when spaces are well-designed. For instance, a vibrant city square with strategically placed benches and green spaces may create a sense of low social density despite a large number of people present.

Social density is particularly interesting when designing interactions in a VE. In the absence of physical constraints, the perception of social density can vary depending on factors such as the size and design of the VE, the number and size of avatars or virtual entities present, and the availability of interactive elements. Guitton et al. present an ethnographic study of merfolk in Second Life, highlighting the challenge of achieving social density in expansive virtual ocean environments [11]. Moore et al. also find that expansive environments in massively multiplayer online games make it challenging to achieve social density and create spaces where it is easy for people to interact [21]. There is limited quantitative work modelling how social density affects interaction in a VE, representing a current gap in how these spaces should be designed for communication and collaboration.

2.3 Focused/Unfocused Interactions

The focus (or lack of focus) during interaction also impacts how people interact and use space. Goffman categorises face-to-face interactions as either *unfocused* or *focused* [9].

Unfocused interactions are characterised by a lack of specific visual attention or shared interest among people. In such scenarios, individuals may coexist in the same space without actively interacting with one another or directing their attention towards a common point of interest. In contrast, focused interactions involve individuals coming together with a shared point of visual attention or a common goal. In these situations, people gather to collaborate, discuss, or participate in activities that require collective focus. Focused interactions often have more defined social regulations, explicit or implicit boundaries for participation, and specific expectations or affordances for involvement. The atmosphere during focused interactions tends to be more structured and purpose-driven. [9].

Focused interactions can be further refined when considering if the rules for social engagement are "tight" or "loose" [7–9]. Social interactions with "tight" rules have highly structured and mutually understood rules for engagement. For example, a presentation to an audience is built on shared and mutually maintained rules around where the focus is and acceptable modes of participation and interruption. In contrast, social interactions with "loose" rules are more unstructured and there can be much flexibility in how the rules, if any, are enforced.

3 Understanding Behaviour in Social VR with High and Low Social Density

The key motivation of our evaluation was to understand how behaviour would be influenced when the virtual environment put constraints on social density within the virtual space. We recruited participants for conversational tasks in an instrumented virtual environment, including a *loosely* and *tightly* structured conversation that would elicit different kinds of behaviours. The virtual environments were sized to create *low* or *high* social density, which we hypothesised would influence how people use space and their comfort in maintaining personal space and turn-taking during conversation.

3.1 Experimental Protocol

For our lab study, we conducted a between-subjects evaluation that compared user experiences within two virtual spaces with different sizes, resulting in either low or high social density as described in Section 3.1.1. We recruited a total of 48 participants, organised into 8 groups of 6 participants each. These groups engaged in conversational tasks as described in Section 3.1.2, with the task order being counterbalanced to mitigate order-related effects. A facilitator was present in the virtual environment to guide participants through the experiment, as described in Section 3.1.3.

To begin, participants were gathered in the same physical space for on-boarding, reading information sheets, and consenting to participate. Participants were assigned unique identifiers: Apple, Banana, Cherry, Dragonfruit, Elderberry, and Fig. The assignment of these identifiers was randomised, each associated with a different physical room to which the facilitators directed each participant. These identifiers were beneficial for both facilitators and participants, aiding in study organisation and participant recognition throughout the evaluation.

As the first task, the Simulator Sickness Questionnaire (SSQ) [17] was administered to assess any effect of the experimental conditions on cybersickness. After completing the SSQ questionnaire, each participant was taken to an individual room, as described in Section 3.2.1. The facilitator gave a short tutorial to each participant, explaining how to fit and control the device, as well as how to enter and navigate the virtual environment; either with joystick movement or teleportation. Participants were given a period of time to accustom themselves to the experience before start the first of two conversation tasks, as described in Section 3.1.2. The order of the conversation tasks was counterbalanced across participant groups. A short questionnaire for Quality of Interaction (QoI), composed of 8 items, was administered after each conversation task, allowing for finer granularity in understanding the impact of each task on the QoI. At the end of the second conversation task, three longer questionnaires were given to the participants to fill out, namely the immersion questionnaire [14], the iGroup Presence Questionnaire (iPQ) [23], and the Social Presence questionnaire [18]. The post-study SSQ was also filled by participants.

To conclude, participants gathered in the same physical room and engaged in a focus group session, during which they shared their perspectives with each other. These encompassed their impressions of the virtual environment, comparative analyses with video calls and face-to-face interactions, their comfort levels, positional dynamics within the virtual space, and their overall experiences in the conversation tasks. Participants were also encouraged to provide feedback on their preferred and least preferred activities, alongside suggestions for potential enhancements. This concluded with thanking the participants for their time and allowing them any questions about the study.

3.1.1 Virtual Environment. Both rooms, as shown in Figure 1, were designed in a circular layout, with one room having a radius of 26.5 meters with low social density, and the other being smaller with a radius of 2.5 meters with high social density. Whyte describes the "effective capacity" of a space in terms of sitting space, where effective capacity is roughly .9 meters of sitting space per person. Within Hall's proxemic zones, this is within the *personal zone* [12]. Moore describes the low social density in one space as a small space where one must "rub elbows with other participants" where it is impossible to watch from a distance [21].

For the 6 participants plus a facilitator, our low social density room allows for 315 meters per person and the high social density room allows for just 2.8 meters per person. The high social density room was designed such that participants were crowded within the *personal and social* proxemic zones, where the low social density provided a large space where participants could move freely. The rooms were intentionally designed to be sparse, with a textured floor plane, smooth walls, and a textured ceiling plane to reduce the impact of environmental elements to influence behaviour. The rooms were circular to provide cornerless spaces after pilots indicated participants often positioned themselves within corners, which introduced arbitrary hot spots of activity.

3.1.2 Conversation Tasks. Goffman describes "tightness" and "looseness" in social settings as the degree to which individuals are expected to be present and attending to interaction, and how strictly social rules must be adhered to [9]. We designed two conversation tasks structured with either "tight" or "loose" social rules, with the hypothesis that different social rules would result in different experiences in a social VR.

Tight Story Task: "Yes, and" Scene in VR. For this task, six participants and one facilitator come together within the VR space to collaboratively construct a story consisting of short statements with tight social rules. The story begins when a participant makes the first statement, for example "I left my house today but I forgot my bicycle helmet." Each following statement must begin with "yes, and" and must be a continuance of the previous statement. Each statement should be short, but must continue the storyline established by the first statement. The facilitator monitored the passage of time and conclude the activity after five minutes. If the story naturally concludes, the facilitator instructed the participants to begin a new story until the overall task lasted five minutes. The story task is structured based on tight social rules that dictate the general content of conversation and how conversation proceeds, where turn-taking will be stilted and formal.

Loose Truths Task: Two Truths and a Lie in VR For this task, six participants and one facilitator come together within the VR space to tell two truths and a lie about themselves with loose social rules. Each participant takes a turn introducing themselves by sharing two truths and one lie about themselves. The group then tries to guess which statement is the lie. Once the lie is discovered, another participant introduces themselves. The facilitator monitored the time and helped move the task along as needed until the overall task lasted five minutes. The truths task is structured based on loose social rules, where individuals may speak informally, interruptions and questions are expected, and turn-taking will be regular and informal.

3.1.3 Facilitator. A facilitator was present in the virtual environment to guide the participants through the conversation tasks. The facilitator began by ensuring all participants were present and not experiencing any technical issues before beginning the conversation task. The facilitator would introduce the task and give participants an opportunity to ask any questions. Once the conversation task began, the facilitator monitored the task timing and facilitated the session as needed to keep the task running for five minutes. The facilitator would conclude each task and provide instructions for the participants to exit VR.

3.2 Experimental Setting and Hardware

3.2.1 Physical Space & Infrastructure. The participants initially gathered in a large room where the experiment began and ended. For the tasks in VR, participants were each taken to a private room that was physically and aurally separated from the other participants. In each private room, there was a desk, chair, and Quest 2 running a customised instance of Mozilla Hubs[1]. Each room was equipped with a 6GHz WiFi router to ensure maximum bandwidth availability and stability. The experiment was run using a modified version of the social VR platform Mozilla Hubs, deployed on a private cloud instance.

The participants used Meta Quest 2 headsets and the facilitator used a Meta Quest Pro headset. All VR users were seated for the duration of the experiment. Participants and facilitator used hand-held controllers for hand movement capture during the experiment.

[1]https://hubs.mozilla.com/

3.2.2 Data Logging. During the experiment, participants' body, hand, and voice data was logged continuously. This was achieved through the use of a custom JavaScript-based subsystem, integrated into a modified version of the Mozilla Hubs client, described in [27]. The data logger collected metrics from the browser environment and directly from the Document Object Model (DOM) which specifies the virtual environment. We augmented the system to collect additional metrics for the purpose of this experiment. Technical details about the data collection is included in the appendices.

Key data streams used for this analysis include the avatar position, direction and rotation based on the headset sensors, position, direction and rotation of each hand based on the hand-held controller sensors, and the amplitude of the participants voice based on the directional microphone within the headset. All this data was updated and saved at a rate, constrained by the frame rate of the user's headset. In order to minimise the number of requests sent to the data collection server, the tick data is buffered and a POST request containing all the data as a JSON-formatted payload is only sent every 4000 ticks. The data collection server is responsible for validating and storing all session data and runs on a separate cloud machine.

3.3 Participants

A total of 48 individuals participated in the experiment. Participants were divided into groups of six for a total of eight sessions, each lasting approximately 1 hour. To ensure a diverse and balanced set of people (i.e., age, gender and nationality), participants were primarily recruited through a professional recruitment company but also from the organisation where the study took place. Since the experiment was conducted in English, all the participants were comfortable in reading, speaking, and writing in English. Additionally, none of them had uncorrected visual, hearing, or motor impairments, and they were instructed to attend the experiment with their glasses or lenses. All participants were aged between 22 and 68 years old with an average age equal to 37.48 and a standard deviation of 12.89. Among all of them, 50% identified as female, 46% as male and the remaining 4% as non-binary (female = 24, male = 22, non-binary = 2). In terms of ethnicity, most of the participants self-identified as being from Europe (40%), followed by Asia (19%), mixed origins (19%) and Africa (2%). 12 participants never had experience with VR before the experiment while 8 of them were experts (e.g., VR designers, or researchers); 22 were novices (e.g., have played less than 5 times) and only 6 were knowledgeable about this technology (e.g., have played 5+ times, or own VR headset). One participant self-identified as a person with a disability (i.e., dyslexia) and one preferred not to say. The majority of participants did not know each other before their involvement in the study. This experimental study was reviewed and approved by an institutional ethics committee.

4 Results

4.1 Quantitative Results

The quantitative results are based on the interaction logs generated by the instrumented virtual environment where we conducted this research. For our 48 participants (excluding the facilitator), our dataset includes 2.42 million logs of 131 minutes of interaction. This data was resampled to 30Hz for the analysis, resulting in 1.48

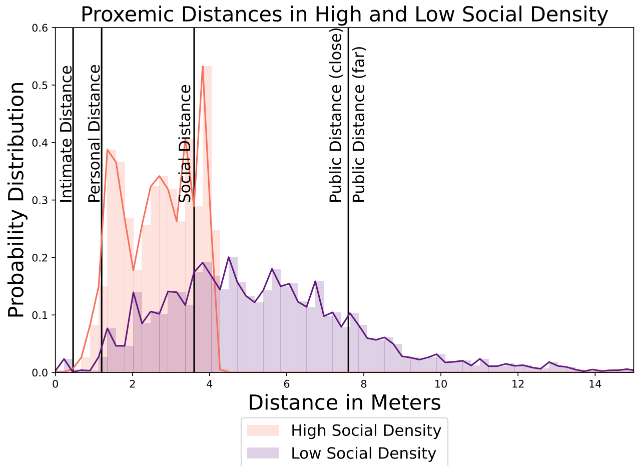
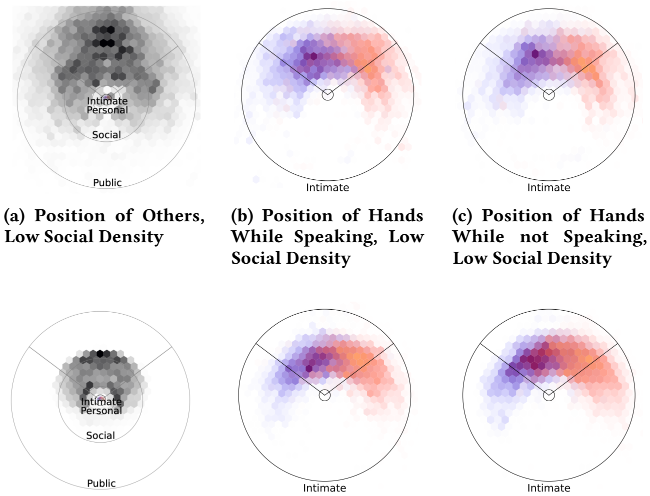


Figure 2: Proxemic analysis compares how people used personal space in the low and high social density conditions.



(a) Position of Others, Low Social Density
(b) Position of Hands While Speaking, Low Social Density
(c) Position of Hands While not Speaking, Low Social Density
(d) Position of Others, High Social Density
(e) Position of Hands While Speaking, High Social Density
(f) Position of Hands While not Speaking, High Social Density

Figure 3: Hexagonal bin plots with proxemic zones overlayed visualise the position of others relative to the field of view for all participants (facilitator excluded). The detailed insets from the Intimate zone visualise the position of the hands while speaking or not speaking. Others are visualised in grey, the left hand is visualised in blue, and the right hand is visualised in orange.

million logs of 131 minutes of interaction. All of the data, scripts, and visualisation techniques used in this analysis will be made publicly available as part of this paper.

Due to an issue with data logging on the facilitator's headset, the facilitator data is unavailable for six of the eight groups. However, quantitative analysis of the social metrics we have developed excludes the facilitator by design for some metrics and these metrics are not affected. Where the facilitator's data is required, we have performed the analysis on the subset of data where the facilitator is included. The subset of the data where the facilitator is present includes over 900,000 interaction logs, which was resampled to over 400,000 logs at 30 frames per second. Given the scale and detail of the data made possible by the instrumented virtual environment, the absence of the facilitator data from some groups has a limited impact on the quantitative analysis.

4.1.1 Social Density: Proximity and Personal Space. Previous works on digital proxemics have been concerned with how people use virtual space and how personal boundaries are maintained in virtual environments. In our design, we specifically put participants in virtual spaces with high and low social densities. In the high social density space, participants could not avoid collisions in the *personal zone* and would have limited space to maintain positions outside the personal zone throughout the task. In the low social density space, participants had ample room to move and could stand at their desired distances within room to move and adjust as needed. Figure 2 shows the proxemic positions for the high and low social density spaces, demonstrating how individuals used personal space in these settings. In line with previous work, we saw that participants avoided collisions in the intimate zone and maintained standing distances in the social and public zones where possible. When constrained in the high social density space, participants maintain social distances, with some collisions in the personal zone. When given space to move, more relaxed proxemics were maintained, with participants spreading across the social and public proxemic zones.

We extended proxemic analysis from previous studies of digital proxemics [4, 12, 15, 28] by adding analysis of how hands aid in defining personal space and how social density might change how people maintain and communicate personal space.

Figure 3 shows a relationship between social density and how people use virtual space and their hands. The position of others relative to the field of view is shown in greyscale, with the proxemic zones shown as concentric rings. In a zoomed view of the intimate zone, the hands are shown in purple (left hand) and orange (right hand) while speaking and not speaking. The facilitator data is excluded from this analysis.

Figure 3 (a) and (d) show results in line with previous research on digital proxemics. In the low social density room, participants tended to stand within the social proxemic zone, with some participants in the personal space, and limited collisions in the intimate space. Participants also tended to keep others within their field of view, and rarely stand with someone directly behind them.

Within the high social density room, there were tight restrictions on personal space and people behaved differently. There are virtually no collisions in the intimate zone, and participants were arranged around the border of the personal zone. Participants has a strong tendency to keep others within their field of view, but the close standing distances forced by the space would make it impossible to see other participants in full. Hotspots are visible directly in front of each participant and to each side, showing the typical configuration in the high social density room to form a circle against the walls.

The hand movements are also affected by social density. Even though in physical reality participants were alone in their room and did not risk hitting anyone or invading their physical space with their hands, participants in the high social density room gestured closer to their own body, as shown in Figure 3 (e) and (f).

Within the low social density room, there was a greater likelihood to ignore intimate collisions, but there was also different use of the hands, in particular while speaking. Gesturing extends further out from the body, sometimes extending beyond the intimate zone. While people often use their hands in interesting ways while speaking [16], looking at position alone does not give a clear picture of how this translates to virtual environments. Hands are also used to claim personal space and add expression while speaking.

4.1.2 Conversation Task: Turning Taking, Silence, and Crosstalk. Participants completed a tightly and loosely structured conversation task, which we designed to elicit stable and unstable conversations for our analysis. For the two groups where the facilitator data is present, we calculated conversational metrics as speaker turn duration, silence duration, and cross-talk duration[1], as shown in Figure 5.

We see differences between the tightly structured story task and the loosely structured truths task. In particular, the challenges for turn-taking can be seen in the increased duration of cross talk during these tasks. The story task was especially challenging in the high social density room, where the conversation involved especially short utterances, long durations of silence, and longer periods of cross talk.

Another metric commonly used in interaction analysis is speaker participation and conversation balance [19]. We visualise this as speaker rank, which describes how dominant speakers were during these conversational tasks. The speaker rank calculates the proportion of time each speaker spent speaking and ranks these in order from most to least for each group, as shown in Figure 4. The facilitator's speech is excluded from this metric.

Figure 4 shows a pattern across all groups, where one speaker often held the floor for roughly half of the time, followed by the next dominant speaker up to a third of the time, with decreasing participation until the least active speaker who made few if any utterances. It was surprising that even in the Truths task, where turning-taking was partially moderated by the facilitator, we see the same pattern of dominance and fall-off in speaker participation. Previous work has looked at influencing speaker participation through visual interventions [19], and similarly found these patterns of behaviour difficult to influence and control.

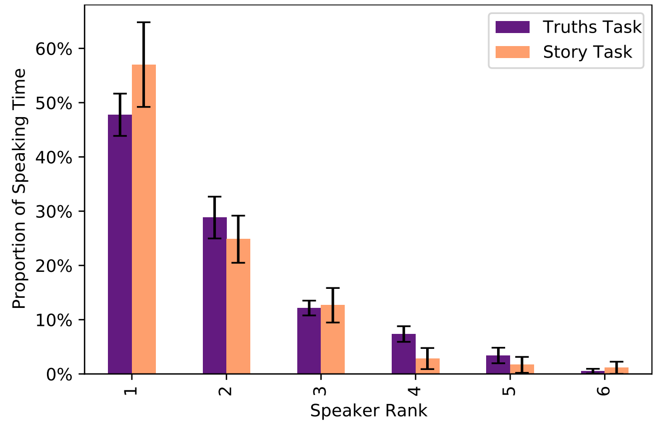


Figure 4: The speaker ranks takes the proportion of time spent speaking for participants in each group and ranks them from most to least time spent speaking. Speaker dominance is very consistent between groups.

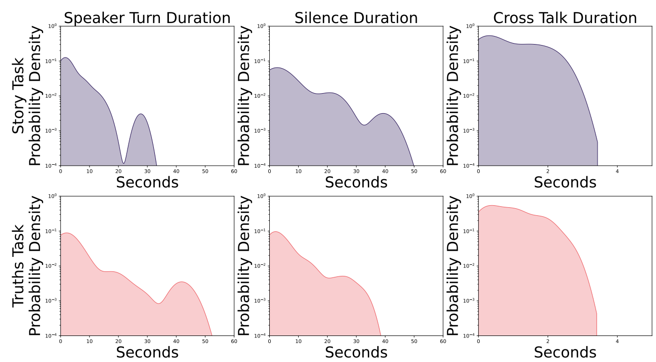


Figure 5: Conversational metrics for the four groups which included the facilitator. Speaker turn duration describes the time a speaker can hold the floor with trailing silence excluded. Silence duration describes the duration of silence when no speaker is active. Cross-talk duration describes the duration when two or more people are speaking simultaneously. Y-axis shows probability density using log scale.

4.2 Questionnaire Results

4.2.1 Quality of Interaction. Questionnaires were administered after each task, allowing to compare the effect of the task on the QoI, as well as increasing the statistical power when comparing the two room configurations. A total of 96 scores were collected per item. Prior to running any comparison, the normality of the scores was tested using the Kolmogorov-Smirnov test. The null hypothesis was rejected ($p < 0.001$); thus, non-parametric tests were used on the scores. Regarding the impact of the task on the QoI, a Wilcoxon test for paired samples revealed no significant statistical difference between the tasks for any of the items on the questionnaire ($p > 0.05$ for all items). We thus focused our analysis on comparing the effect of the different room configurations on the items. Figure 6 depicts the violin plot of the score distribution for every item in the QoI questionnaire. Please note that the scales were inverted, when necessary, to ease readability; in all cases, higher scores signify positive responses. The plot shows the distribution of the scores, along with the median value, separately for the low social density configuration (left, darker colour) and the high social density configuration (right, lighter colour). In general, we can observe high values for almost all the items in the questionnaire, with median values between 5 and 7; the notable exception is for the item "Involvement", for which results are fairly spread, with a median value of 4. This might be due to the phrasing of the question ("How much did your experience in the virtual environment seem consistent with your real-world ones?"), which

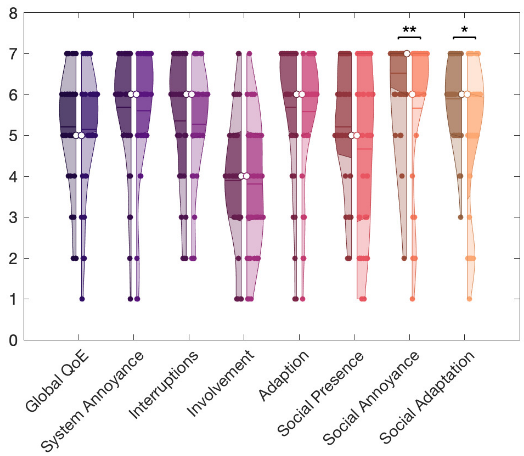


Figure 6: Violin plot showing the results of the QoI questionnaires (higher values are better). Left and right sides refer to the low density and high density room configurations, respectively. Statistical significant difference, when present, is indicated with asterisks on top (*: $p < 0.05$; **: $p < 0.01$; ***: $p < 0.001$).

led to confusion among the participants. A Mann-Whitney test for unpaired samples performed on the items reveals no statistical difference for 6 out of the 8 items in the questionnaire (Global QoE: $Z = -0.061, p = 0.952, r = 0.006$; System Annoyance: $Z = 0.023, p = 0.982, r = 0.002$; Interruptions: $Z = -0.525, p = 0.599, r = 0.054$; Involvement: $Z = 0.336, p = 0.737, r = 0.034$; Adaption: $Z = 0.890, p = 0.374, r = 0.091$; Social Presence: $Z = -1.131, p = 0.258, r = 0.115$), whereas the difference was found to be significant for Social Annoyance ($Z = -2.621, p = 0.009, r = 0.268$) and Social Adaptation ($Z = -2.320, p = 0.020, r = 0.237$), with small to medium effect sizes in both cases. Results seem to indicate that the high social density room configuration led to disruption in terms of listening to other participants (Social annoyance: "I was able to understand my partners' talking") and cooperating with them (Social Adaptation: "My partners and I worked together well"), with respect to the low social density room configuration, which offered more space to move and a more comfortable setting.

4.2.2 Presence, Immersion, Social Presence. Questionnaires for Immersion, Presence, and Social Presence were administered after completing both tasks, leading to 48 scores per item. In all cases, we grouped the items according to the original reference, and we compared the distributions associated with the high and low density room configurations.

Figure 7a displays the distribution of the scores associated with the factors "Immersion" and "Person-Virtual Environment (VE) Interaction". The values were averaged to ease comparison between different factors, and ordered so that higher scores indicate a positive outcome. We can see that, for both room configurations, scores are favorable for both factors, indicating high immersion levels. A normality check conducted using the Kolmogorov-Smirnov test rejected the null hypothesis; thus, we employed non-parametric statistical tests. A Mann-Whitney U-test revealed no significant differences between the room configurations for Immersion ($Z = 0.8418, p = 0.3999, r = 0.1215$) nor for Person-VE Interaction ($Z = -0.2100, p = 0.8336, r = 0.0303$).

Figure 7b depicts the distribution of the scores gathered using the Social Presence questionnaire, aggregated in the factors "Quality of Interaction", "Presence/Immersion", and "Social Meaning". Please note that for this questionnaire, the Likert scale ranged from 1-5. As with the other questionnaires, items were averaged per factor and scaled so that higher values would indicate positive outcomes. We can see that high values were reported for factors "Presence/Immersion" and "Social Meaning", corroborating the findings of the previous questionnaires. Values associated with "Quality of Interaction" sit slightly lower. The low scores stem mainly from the answers given to the first two questions, which mapped more to an emotional connection: "I was able to feel my partners' emotions during the experience", and "I was sure that my partners often felt my emotion". It is worth noting that the Social Presence questionnaire was originally formulated to be used with pairs that knew each other, and validated with a task (namely, photo sharing) which allows for an emotional connection between people. Conversely, our tasks were not designed to foster an emotional bonding between participants, which explains why low scores were given to the corresponding questions. Test using Kolmogorov-Smirnov rejected the normality hypothesis, and the Mann-Whitney U-test applied to the data failed to reject the null hypothesis for all factors (Quality of Interaction: $Z = 0.1497; p = 0.8810; r = 0.0218$; Presence/Immersion: $Z = -0.4051; p = 0.6854; r = 0.0591$; Social Meaning: $Z = -1.4946; p = 0.1350; r = 0.2180$).

Finally, figure 7c shows the violin plot of the scores associated with the items in the iGroup Presence questionnaire, grouped in the factors "Presence", "Social Presence", "Involvement", and "Experienced Realism". As for the previous questionnaire items, scores within each factor were averaged to facilitate comparison, and they were flipped if necessary to always map higher scores to a more positive outcome. We can see high values for all the factors in the questionnaire, except for the "Experienced Realism", which received lower scores on average (median value $M = 3$ and $M = 3.25$ for the small and big room configuration, respectively). The low scores associated with realism can be explained by considering that our environment is not a faithful replica of the real world, instead being designed as rather barebone and simplified. Thus, it is reasonable that users gave low scores to questions such as "The virtual world seemed more realistic than the real world", and "How real did the virtual world seem to you". As for the previous case, we ran a normality test using Kolmogorov-Smirnov, which rejected the hypothesis of normality. The Mann-Whitney U-test we applied to the data failed to unveil any significant difference between the room configurations, for all factors (Presence: $Z = -0.4239; p = 0.6716; r = 0.0618$;

Social Presence: $Z = -1.6233; p = 0.1045; r = 0.2368$; Involvement: $Z = -0.3312; p = 0.7405; r = 0.0483$; Experienced Realism: $Z = -0.9726; p = 0.3307; r = 0.1419$).

4.2.3 Simulator Sickness. The SSQ was administered before and after the VR experiment, as suggested by the literature [5]. For each participant, three factors, namely Nausea N, Oculomotor disturbance O, and Disorientation D, were computed, according to the original reference, and the total score was computed by summing the three factors and applying a weight of 3.74 [17]. The original paper indicates the following ranges for interpreting the scores: negligible (< 5), minimal (5 − 10), significant (10 − 15), and concerning (15 − 20). In general, we witness a marked increase in symptoms after the VR test. The reported results for Nausea are negligible before the test ($N = 4.56$), but they turn significant after the test ($N = 10.78$); for Oculomotor disturbance, they are minimal before the test ($O = 8.73$), and they become significant after ($O = 13.18$); similarly, for Disorientation, values are significant before the test ($D = 10.29$), but they become concerning after ($D = 19.06$). The total score before the test is equal to 88.21, whereas, after the test, it increases to 160.94. The difference between symptoms before and after the experiment is significant, according to a Wilcoxon signed rank test, only for Nausea ($Z = -2.2515, p = 0.0244, r = 0.2373$), but not for Oculomotor disturbance ($Z = -0.9613, p = 0.3364, r = 0.1013$) or Disorientation ($Z = -1.6820, p = 0.0926, r = 0.1773$), and neither for the total score ($Z = -1.8705, p = 0.0614, r = 0.1972$).

We also test for statistical differences between the room configurations, to see whether they had an effect on the sickness symptoms, using a Mann-Whitney U-test. Results indicate no significant differences in Nausea ($Z = 0.5891; p = 0.5558; r = 0.0621$), Oculomotor disturbance ($Z = 1.8261, p = 0.0678, r = 0.1925$), or Disorientation ($Z = -0.9189, p = 0.3581, r = 0.0969$), nor for the total score ($Z = 0.4423, p = 0.6583, r = 0.0466$).

4.3 Qualitative Results

The qualitative results were analysed from the focus group transcripts using a three stage coding process. The transcript data includes 155 minutes of discussion with the 48 participants in groups of 6. Participant identifiers are shown as the group number and participant ID.

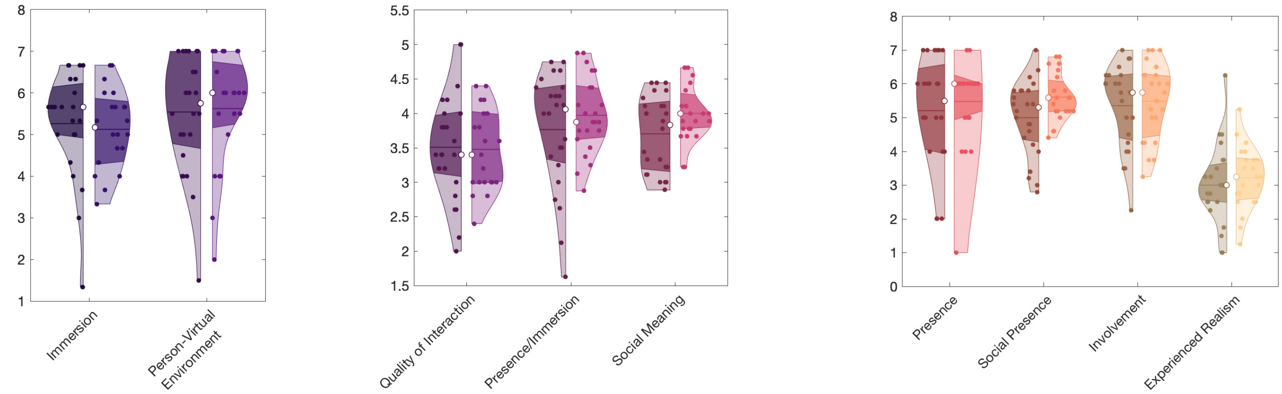
4.3.1 Social Density: Visceral Reactions to Personal Space. Participants had visceral reactions to personal space and how this impacted their experiences in the high or low social density spaces. By design, the high social density space gave each participant limited personal space. The space was described as "too small," "crowded," and even "claustrophobic." Participants in the high social density space were concerned about being too close to others, struggled to find space for themselves, and described the sensation of constantly wanting to step back. The sense of feeling constrained in the high social density space is also visible in the hand tracking data shown in Figure 3, where participants gestured closer into their own bodies even though their physical space was not constrained.

The option to use space more freely in the low social density space gave participants more ways of expressing themselves in the space. For example, Participant 3D stated that "I kind of move closer when accusing someone just to see how they react" when describing the Truths conversation task. Participant 5C and 7B commented on how moving towards others while they were speaking helped them to better hear the speaker. The ability to use space freely in the low social density space afforded these expressive and playful actions, which were not possible in the uncomfortable close distances of the high social density space.

Participants described how challenging it could be to keep others within their field of view in the high social density space, where the proportion of the body that is visible is a well established issue in face-to-face interactions [6, 12]. In the high social density space, participants would be standing within the *personal* proxemic zone, preventing visibility of complete avatars or the whole group simultaneously. Participant 6A noted particular difficulties in the high social density space, stating that "it was too small. I tried to look at people... but there was always someone in the way." In contrast, in the low social density space, participants benefited from arranging themselves specifically so they could see others. When given freedom to stand at one's desired distance relative to others, participants spread comfortably across the personal and social zones, as shown in Figure 3. This enabled better visibility of other participants, and resulted in a better quality of interaction.

4.3.2 Conversation: Who's Line is it Anyway? Conversation in VR presented specific challenges for turn taking and demonstrating attention. The importance of non-verbal cues in social VR has been explored thoroughly in related research [1, 2, 20], with different capabilities for non-verbal communication. In our evaluation, non-verbal communication was limited to head, hand, and body movements and participants were specifically concerned about the challenges of turn-taking during the conversation tasks. Missing social cues influenced the ebb and flow of conversations. Participant 2C highlighted, "you don't want to interrupt somebody because you cannot see who's speaking." 7B revealed, "I kept on thinking who's gonna talk next." The challenge of uncertain timing and potential interruptions was further exemplified by participant 2B, "Felt like there was a bit too much silence. I was like, should I say something or not?" The anxiety of taking turns and the impulse to stay engaged were encapsulated by participant 7B, "Kept on thinking who's gonna talk next." The struggle of identifying gaps in the conversation to make a contribution was highlighted by participant 2D, "Don't want to interrupt somebody because you cannot see who wants to start speaking." The issue of timing and breaks in communication was evident as participant 8A mentioned, "Unnatural breaks... people are waiting to start speaking but nobody wants to start."

A key challenge to conversational turn-taking was displaying and interpreting attention while listening and speaking. 6E shared, "I was not sure if people are listening, I think it's a lack of eye contact, you don't know if other people are looking at you or not." Another participant (2B) underlined, "In the real world, you can see who's coming to talk." The simple signal of head gaze was effective for some participants. Participant 1D stated that "when somebody is speaking, you the people are really looking and giving attention." Participant 2B stated that "I did feel like people were listening to me because all the heads were coming to me." Although head gaze contains less signal than eye gaze when demonstrating attention



(a) Immersion questionnaire (b) Social presence questionnaire (c) iGroup presence questionnaire

Figure 7: Violin plot showing the results of the immersion, social presence, and iGroup presence questionnaires, aggregated per item (higher values are better). The left and right sides refer to the low density and high density room configurations, respectively.

[25], this was useful for participants in understanding if others were paying attention.

5 Conclusion

This paper explores social density as a factor in social interactions in a social immersive environment. Our between subjects study explored social factors like proximity, conversation, and quality of interaction in a high and low social density space, demonstrating how people interact differently across these kinds of spaces. This mixed methods approach combines quantitative and qualitive methods to provides new insights into how interaction unfolds in social immersive environments.

Acknowledgments

This work was supported through the European Union Horizon Europe research and innovation programme, under grant agreement No 101070109 (TRANSMIXR) and the ERC Consolidator Grant FUSION, proposal 101126024, (funded by the UK Horizon guarantee scheme EPSRC project EP/Z000432/1).

All of the code, data, and tools for analysis used in this paper are openly available on GitHub: https://github.com/julierthanjulie/social-density-xr

References

[1] A. Abdullah, J. Kolkmeier, V. Lo, and M. Neff. Videoconference and embodied vr: Communication patterns across task and medium. *Proceedings of the ACM on Human-Computer Interaction*, 5(CSCW2):1–29, 2021.
[2] N. Aburumman, M. Gillies, J. A. Ward, and A. F. d. C. Hamilton. Nonverbal communication in virtual reality: Nodding as a social signal in virtual interactions. *International Journal of Human-Computer Studies*, 164:102819, 2022.
[3] A. Baum and S. Koman. Differential response to anticipated crowding: Psychological effects of social and spatial density. *Journal of Personality and Social Psychology*, 34(3):526, 1976.
[4] S. Benford, D. Snowdon, A. Colebourne, J. O'Brien, and T. Rodden. Informing the design of collaborative virtual environments. In *Proceedings of the 1997 ACM International Conference on Supporting Group Work*, GROUP '97, pp. 71–80. Association for Computing Machinery, New York, NY, USA, Nov. 1997. doi: 10.1145/266838.266866
[5] P. Bimberg, T. Weissker, and A. Kulik. On the usage of the simulator sickness questionnaire for virtual reality research. In *2020 IEEE conference on virtual reality and 3D user interfaces abstracts and workshops (VRW)*, pp. 464–467. IEEE, 2020.
[6] J. Gehl. *Life between buildings: using public space*. The Danish Architectural Press, 2001.
[7] M. Gelfand. *Rule makers, rule breakers: Tight and loose cultures and the secret signals that direct our lives*. Scribner, 2019.
[8] M. J. Gelfand, J. L. Raver, L. Nishii, L. M. Leslie, J. Lun, B. C. Lim, L. Duan, A. Almaliach, S. Ang, J. Arnadottir, Z. Aycan, K. Boehnke, P. Boski, R. Cabecinhas, D. Chan, J. Chhokar, A. D'Amato, M. Subirats, I. C. Fischlmayr, R. Fischer, M. Fülöp, J. Georgas, E. S. Kashima, Y. Kashima, K. Kim, A. Lempereur, P. Marquez, R. Othman, B. Overlaet, P. Panagiotopoulou, K. Peltzer, L. R. Perez-Florizno, L. Ponomarenko, A. Realo, V. Schei, M. Schmitt, P. B. Smith, N. Soomro, E. Szabo, N. Taveesin, M. Toyama, E. V. de Vliert, N. Vohra, C. Ward, and S. Yamaguchi. Differences between tight and loose cultures: A 33-nation study. *Science*, 332(6033):1100–1104, 2011. doi: 10.1126/science.1197754
[9] E. Goffman. *Behaviour in Public Places: Notes on the Social Organization of Gatherings*. Free Press of Glencoe, NY, USA, 1963.
[10] S. W. A. Group. *From lizarding to lingering: how we really behave in public spaces*. The Guardian, 2020.
[11] M. J. Guitton. Swimming with mermaids: Communication and social density in the Second Life merfolk community. *Computers in Human Behavior*, 48:226–235, 2015. doi: 10.1016/j.chb.2015.02.004
[12] E. T. Hall. The Hidden Dimension: man's use of space in public and private The Bodley Head. *London, Sydney, Toronto*, 121, 1969.
[13] B. Hillier and J. Hanson. *The social logic of space*. 1988. doi: 10.4324/9780429450174-9
[14] S. Hudson, S. Matson-Barkat, N. Pallamin, and G. Jegou. With or without you? interaction and immersion in a virtual reality experience. *Journal of Business Research*, 100:459–468, 2019.
[15] T. Iachini, Y. Coello, F. Frassinetti, and G. Ruggiero. Body space in social interactions: a comparison of reaching and comfort distance in immersive virtual reality. *PloS one*, 9(11):e111511, 2014.
[16] A. Kendon. *Conducting interaction: Patterns of behavior in focused encounters*, vol. 7. CUP Archive, 1990.
[17] R. S. Kennedy, N. E. Lane, K. S. Berbaum, and M. G. Lilienthal. Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness. *The international journal of aviation psychology*, 3(3):203–220, 1993.
[18] J. Li, Y. Kong, T. Röggla, F. De Simone, S. Ananthanarayan, H. De Ridder, A. El Ali, and P. Cesar. Measuring and understanding photo sharing experiences in social virtual reality. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, pp. 1–14. Association for Computing Machinery, New York, NY, USA, 2019.
[19] J. V. Li, M. Kreminski, S. M. Fernandes, A. Osborne, J. McVeigh-Schultz, and K. Isbister. Conversation balance: A shared vr visualization to support turn-taking in meetings. In *Extended Abstracts of the 2022 CHI Conference on Human Factors in Computing Systems*, CHI EA '22. Association for Computing Machinery, New York, NY, USA, 2022. doi: 10.1145/3491101.3519879

[20] D. Maloney, G. Freeman, and D. Y. Wohn. " talking without a voice" understanding non-verbal communication in social virtual reality. *Proceedings of the ACM on Human-Computer Interaction*, 4(CSCW2):1–25, 2020.
[21] R. Moore, E. Hankinson Gathman, and N. Ducheneaut. From 3D space to third place: The social life of small virtual spaces. *Human Organization*, 68(2):230–240, 2009. ISBN: 0018-7259 Publisher: Society for Applied Anthropology.
[22] R. Oldenburg. *The great good place: Cafes, coffee shops, bookstores, bars, hair salons, and other hangouts at the heart of a community.* Da Capo Press, 1999.
[23] T. W. Schubert. The sense of presence in virtual environments: A three-component scale measuring spatial presence, involvement, and realness. *Z. für Medienpsychologie*, 15(2):69–71, 2003.
[24] D. Thery and B. F. Katz. Auditory perception stability evaluation comparing binaural and loudspeaker Ambisonic presentations of dynamic virtual concert auralizations. *The Journal of the Acoustical Society of America*, 149(1):246–258, 2021. ISBN: 0001-4966 Publisher: Acoustical Society of America.
[25] O. Špakov, H. Istance, K.-J. Räihä, T. Viitanen, and H. Siirtola. Eye gaze and head gaze in collaborative games. In *Proceedings of the 11th ACM Symposium on Eye Tracking Research & Applications*, ETRA '19. Association for Computing Machinery, New York, NY, USA, 2019. doi: 10.1145/3317959.3321489
[26] W. H. Whyte. The social life of small urban spaces. *Publisher: Conservation Foundation Washington, DC*, 1:466–468, 1980.
[27] J. Williamson, J. Li, V. Vinayagamoorthy, D. A. Shamma, and P. Cesar. Proxemics and Social Interactions in an Instrumented Virtual Reality Workshop. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, CHI '21, pp. 1–13. Association for Computing Machinery, New York, NY, USA, May 2021. doi: 10.1145/3411764.3445729
[28] J. R. Williamson, J. O'Hagan, J. A. Guerra-Gomez, J. H. Williamson, P. Cesar, and D. A. Shamma. Digital proxemics: Designing social and collaborative interaction in virtual environments. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, pp. 1–12. Association for Computing Machinery, New York, NY, USA, 2022.

A Data Collection

A cloud server with a single HTTP POST endpoint was set up for collating all the user data gathered during an experiment session. The server is implemented using the Go programming language to achieve adequate performance and keep system load to a minimum. Running as a background process, the server listens for incoming POST requests on TCP port 6000. Further, all requests are handed to the sever through a Nginx reverse proxy, which also takes care of CORS policy validation to allow the Mozilla Hubs clients to communicate directly with the data collection server via AJAX.

Upon reception of a request on the right endpoint, a streaming JSON decoder is instantiated, ready to receive the payload body of the POST request. Once the entire payload is received and validated, the program checks for the presence of all required fields. If all required fields are present, the decoded payload is converted to a comma-separated format and appended to a compressed CSV file using a streaming GZip compressor. The server also adds a UNIX timestamp to each record, which can be used correct possible time drift and/or inaccuracies in the timestamps received from the clients. To prevent file corruption through concurrent access, the write operation is guarded by a mutex. After a successful write, the request handler returns a message with HTTP status 200 to the client.

If the submitted data did not pass validation, the server returns an error with HTTP status 400 to the client. If the data could not be written to the file, an error with HTTP status 500 is returned.

Through the use of GZip compression, the data collected during a typical session, which amounts to about 2 GB, can be compressed to about 500 MB, thus keeping storage space use to a minimum. Further, through the use of a streaming compressor, the file handle can be held onto without having to close and reopen the file for every request.

B Metrics Gathered from Browser Environment

`timestamp`	Device's UNIX timestamp in milliseconds
`fps`	Current frame rate
`uuid`	UUID
`user_agent`	Device user agent
`isBrowser`	Device type
`isLandscape`	Device orientation
`isWebXRAvailable`	VR availability
`avatarID`	Avatar ID
`isHeadsetConnected`	Headset connection status
`isRecording`	Recording status
`pathname`	Current URL
`urlQuery`	Query section of the URL

C Metrics Gathered from DOM Tree

`isLoaded`	Has user finished loading
`isEntered`	Has user joined room
`isFlying`	Is user flying
`isVisible`	Is user visible
`isSpeaking`	Is user speaking
`isMuted`	Is user muted
`volume`	Current user volume
`rigPosX, rigPosY, rigPosZ`	Avatar position (X, Y, Z)
`rigDirectionX, rigDirectionY, rigDirectionZ`	Avatar direction (X, Y, Z)
`rigQuatX, rigQuatY, rigQuatZ, rigQuatW`	Avatar quaternion rotation (X, Y, Z, W)
`povPosX, povPosY, povPos`	POV position (X, Y, Z)
`povDirectionX, povDirectionY, povDirectionZ`	POW direction (X, Y, Z)
`povQuatX, povQuatY, povQuatZ, povQuatW`	POV quaternion rotation (X, Y, Z, W)