Socially Late, Virtually Present: The Effects of Transforming Asynchronous Social Interactions in Virtual Reality

Portia Wang
Stanford University
Stanford, California, USA
portiaw@stanford.edu

Mark R. Miller
Illinois Institute of Technology
Chicago, Illinois, USA
mmiller30@iit.edu

Anna C.M. Queiroz
Stanford University
Stanford, California, USA
acmq@stanford.edu

Jeremy N. Bailenson
Stanford University
Stanford, California, USA
bailenson@stanford.edu

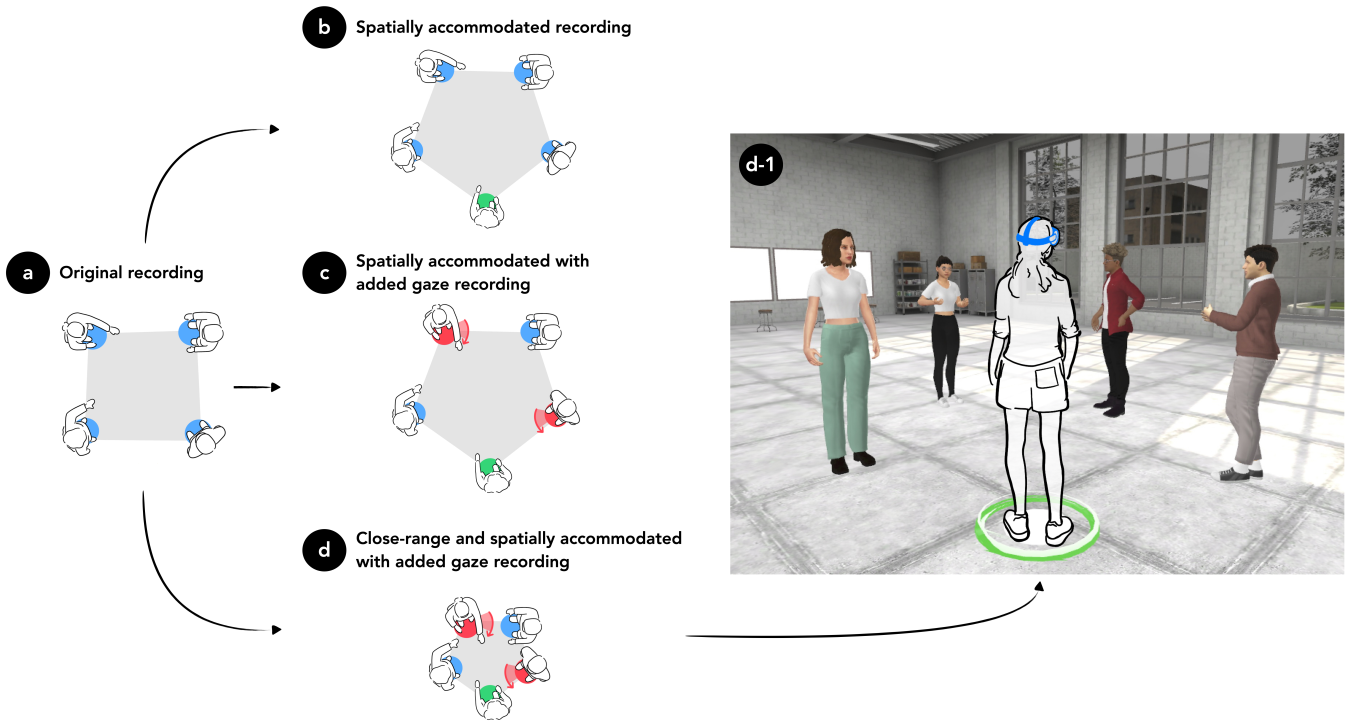

Figure 1: Transforming asynchronous VR social interactions. In (a–d), the green circle denotes the position of the asynchronous user, the blue circles denote the positions of the recorded avatars, and the red circles denote positions of recorded avatars with added gaze. (a) Top-down schematic of the recorded four-person discussion. (b) Top-down schematic of the spatially accommodated recording. (c) Top-down schematic of the spatially accommodated with added gaze recording. (d) Top-down schematic of the close-range and spatially accommodated with added gaze recording. (d-1) An asynchronous user joining a transformed group discussion.

ABSTRACT

Social Virtual Reality (VR) typically entails users interacting in real time. However, asynchronous Social VR presents the possibility of combining the convenience of asynchronous communication with the high presence of VR. Because the tools to easily record and replay VR social interactions are fairly new, scholars have not yet examined how users perceive asynchronous VR social interactions,

CHI '24, May 11–16, 2024, Honolulu, HI, USA

ACM ISBN 979-8-4007-0330-0/24/05
https://doi.org/10.1145/3613904.3642244

and how nonverbal transformations of recorded interactions influence user behavior. In this work, we study nonverbal transformations of group interactions around proxemics and gaze and present results from an exploratory user study (N=128) investigating their effects. We found that the combination of spatial accommodation and added gaze increases social presence, perceived attention, and mutual gaze. Results also showed an inverse relationship between interpersonal distance and perceived levels of dominance and threat of the recorded group. Finally, we outline implications for educators and virtual meeting organizers to incorporate these transformations into real-world scenarios.

CCS CONCEPTS

• **Human-centered computing** → User studies; **Virtual reality**.

KEYWORDS

Virtual Reality, Transformed Social Interaction, Social Interaction, Proxemics, Eye Gaze

ACM Reference Format:
Portia Wang, Mark R. Miller, Anna C.M. Queiroz, and Jeremy N. Bailenson. 2024. Socially Late, Virtually Present: The Effects of Transforming Asynchronous Social Interactions in Virtual Reality. In *Proceedings of the CHI Conference on Human Factors in Computing Systems (CHI '24), May 11–16, 2024, Honolulu, HI, USA.* ACM, New York, NY, USA, 19 pages. https://doi.org/10.1145/3613904.3642244

1 INTRODUCTION

Virtual Reality (VR) enables social communication beyond what is possible in traditional media. Unlike video conferencing, VR allows users to engage in more realistic social interactions by allowing them to communicate nonverbally and engage with social scenes through natural head and hand movement. In particular, the incorporation of nonverbal cues such as gaze, hand gestures, and proxemics (i.e., interpersonal distance behavior) is crucial in creating a sense of presence in synchronous social VR settings [14, 21, 55, 77]. Presence, the subjective sense of "being there", differs from immersion, the objective technical capability of the VR system [75, 92]. Instead of having cameras in fixed locations, VR allows users to view immersive content at novel viewpoints through rotating and moving their heads and bodies, consequently heightening their sense of presence [16, 93]. As the commercial use of VR headsets (e.g., Meta Quest 2 and Pico 4) [1] and social VR platforms (e.g., VRChat, ENGAGE, Horizon Worlds) becomes increasingly prevalent, it is ever more important to consider the users' sense of presence as they engage in immersive social interactions.

In this work, we explore the topic of *asynchronous social interactions*, the notion of participating and observing past social interactions. Asynchronous communication has become an integral part of our daily lives, encompassing mediums such as emails, text messages and content through video sharing software such as YouTube and TikTok. Its importance lies in its ability to bridge temporal and spatial gaps, enabling individuals to engage with content and social conversations at locations and times that are convenient for them. Because of this, we see potentials for combining the convenience of asynchronous communication with the high presence of VR, effectively extending social VR from its typical real-time format into an asynchronous one.

Consider the following scenario. A student missed out on group discussions and lectures due to time zone differences. To catch up, they participate in the discussions and lectures in asynchronous social VR. Using it, the student immersively engages with the recorded social interactions as if they were occurring in real time, allowing them to make up for the missed content despite the temporal constraints. However, it is not clear how we can design efficient and engaging asynchronous social VR. We do not fully understand how to seamlessly integrate new users into previously recorded and reconstructed social scenes. Simply placing a new user into such environments can lead to the feeling of being a fly-on-the-wall, as opposed to an assimilated member of the social group. Following the same scenario, a new student observing a recorded group discussion, having missed the live discussion, will likely find themselves spatially excluded from the recorded group and receive little attention from the members of the group. This naive approach of reconstructing past social interactions could therefore undermine the benefits of social VR. We therefore argue that it is important to explore how we can transform recorded group interactions to acclimate new users into past social scenes. At the same time, given the importance of nonverbal cues in facilitating real-time immersive interactions [14, 21, 55, 77], we argue that it is critical to design and study transformations of nonverbal behavior in asynchronous social VR.

In this paper, we focus on nonverbal transformations of recorded group social interactions in VR. Through investigating how new users perceive these transformed social interactions and further developing guidelines for mitigating potential risks (e.g., ethical and privacy concerns) and improving the new user's subjective experiences, we envision these transformations playing a vital role in reshaping the way we connect with others across time and space. While past works in asynchronous social interactions in VR have compared time-machine-like VR experience sharing with co-watching 360-degree videos on desktop and in VR [83], explored asynchronous collaboration [15, 19], learning [13], preservation of causality for co-dependent events [23], and reviewing of past single-user experiences [54], there is little work studying how we can transform recorded social scenes to acclimate new users into past social interactions and their implications.

In light of this, we design and study nonverbally-transformed asynchronous social VR. We extended previous methods for manipulating proxemics and gaze by sampling and altering the rich tracking data of group discussions collected through the social VR platform ENGAGE. In an exploratory user study with 128 participants, we compared four conditions that transformed recorded discussions in different ways, namely (1) unmodified, (2) spatially accommodated at the same interpersonal distance, (3) spatially accommodated at the same interpersonal distance with added gaze, and (4) spatially accommodated at a closer interpersonal distance with added gaze. Notably, our results on the self-reported user experience and behavioral measurements revealed that the combination of spatial accommodation and inserting gaze increases social presence, perceived attention, and amount of mutual gaze. The results also showed an inverse relationship between interpersonal distance and perceived threat, where interpersonal distance is affected by both the proxemics transformations and participants' deliberate movement in the physical space. Based on our findings, we present practical implications for real-world asynchronous scenarios. Specifically, when facilitating asynchronous VR social interactions, we recommend practitioners applying both proxemics and gaze transformations by referencing the original recordings for gaze patterns and interpersonal distance. Additionally, we recommend employing mechanisms for directing new user attention towards the periphery or placing important avatars near the center of the user's viewing perspective. Finally, we recommend supplementing nonverbal transformations with techniques optimizing for other objectives such as memory recall.

Taken together, our contributions are threefold. First, we extend the preliminary approaches developed by Wang et al. [84] for transforming VR tracking data around proxemics and gaze to assimilate new users into recorded social interactions. Then, we discuss self-reported responses and behavioral findings from a between-participant user study with four conditions varying in the amount of proxemics and gaze manipulations, where each participant observed two recorded group discussions with guiding prompts. Finally, given our findings, we outline design and ethical implications for practitioners such as educators and virtual meeting organizers to integrate these techniques into scenarios beyond the evaluated use case.

2 RELATED WORK

In this section, we discuss avenues of research relevant to ours and our contributions to each. We first provide an overview of works studying nonverbal communication in virtual social interactions. We then describe research on transforming social interactions in VR, and finally related works on perceiving virtual group social interactions.

2.1 Nonverbal Behavior in Virtual Social Interactions

Immersive technologies such as Virtual and Augmented Reality are capable of tracking user motion (e.g., head and hand movement) and mapping them onto digital avatars to facilitate social interactions across multiple users. Leveraging this exact affordance, many have studied how users interact and perceive others in a virtual environment, for instance group formation and proxemics (i.e., the study of interpersonal distance and personal space). Williamson et al. [86] studied how participants in an academic workshop interacted in virtual spaces using VE Mozilla Hubs. During the workshop programme, participants socialized in a large outdoor environment as well as breakout rooms for smaller group activities. The results from 26 participants, 3 of whom joined in VR and 23 through web browsers, showed that smaller spaces facilitated formations of more cohesive groups while larger spaces enabled more flexible personal spaces yet limited formation of smaller groups. Relatedly, Huang et al. [38] investigated personal space in Augmented Reality (AR) between an agent (i.e., a computer-controlled virtual being) and a human user using the framework of proxemics theory. The results showed that users kept a farther distance from male agents than they did from female agents, validating previous works studying gender-based effects in proxemics [34, 85]. The authors also observed elevated skin conductance responses for some of the virtual agents and found that users maintained a closer distance from human-like virtual agents than from a pillar.

One theory encapsulating both the concepts of personal space and gaze is that of Argyle and Dean's [3] equilibrium theory. The theory posits that individuals seek to maintain an equilibrium level of intimacy, which can be influenced by factors such as physical proximity and eye contact. The theory further predicts that a closer interpersonal distance between two individuals, which increases the level of intimacy, will yield less eye contact, thereby decreasing the level of intimacy back to the equilibrium level. Conversely, a higher interpersonal distance is predicted to yield more eye contact to compensate for the drop in intimacy resulting from the greater physical separation. Bailenson et al. [9] examined the relationship between interpersonal distance and mutual gaze in immersive virtual environments by having participants interact with a virtual human who is standing still and approaching the participants. The authors found that the interpersonal distance between participants and the virtual avatar was greater when they engaged in mutual gaze with the virtual avatar. Their results also showed that participants moved farther away when the virtual humans invaded their personal space. Other works also corroborated the need to maintain an appropriate level of intimacy in dyadic or multi-agent VR interactions [9, 53], virtual classroom and conference settings [56, 86] as well as non-immersive virtual communities such as Second Life [26, 91].

We extend existing work on equilibrium theory by examining how transformations of asynchronous group social interactions in immersive virtual environments influence the nonverbal behavior and self-reported measurements of users. When shown recorded group discussions in VR, do users still seek to maintain the level of intimacy as they do when interacting with others in real-time in virtual environments?

2.2 Transforming Social Interactions in Virtual Reality

Bailenson et al. [6] proposed the concept of Transformed Social Interaction (TSI) as techniques that alter the nature of social interactions through decoupling signals such as nonverbal behavior and transforming them before rendering these signals to users in virtual environments. The authors highlighted how digital modifications of social interactions can surpass the constraints observed in face-to-face interactions. Later works have studied modifications of avatar self-representations, for example how embodiment of an elderly person in a perspective taking task influences intergenerational attitudes [62], and how embodying the same uniform avatar within a group differs from wearing an avatar that resemble the users themselves [32].

Scholars have also leveraged TSI to study the effects of transforming nonverbal behavior by manipulating synchrony [10, 67, 68, 81], and inserting gaze towards the user [2, 7, 30, 46, 67, 69]. Bailenson and Yee [10] studied mimicry in VR by having a virtual avatar mimic the head movement of a user at a 4-second delay and read a persuasive message. The authors found that the mimicking agents, in comparison to the nonmicking agents, were more persuasive and perceived by the study participants to have more positive traits.

In a study conducted by Tarr et al. [81], the authors updated the movements of characters participating in a joint movement activity to manipulate synchrony and found that participants in the high synchrony condition reported higher social closeness. Bailenson et al. [7] proposed a transformation for augmenting gaze by rendering interactants' head orientations such that they could simultaneously look at multiple users concurrently while delivering a persuasive message. They found that compared to other gaze conditions, women agreed more to the persuasive message in the augmented gaze condition. Despite not finding evidence that augmented gaze facilitated memory recall, the authors highlighted the potential of applying such transformations in contexts such as remote learning and advertising given their ability to impart personalizable nonverbal attention towards VR users.

Our work builds on previous TSI work by proposing methods for transforming the proxemics and head orientations (i.e., a proxy for gaze) of recorded group discussions in VR. Unlike previous research, we focus on evaluating the effects of these transformations on asynchronous social interactions in immersive virtual environments, where a new user is immersed in past recorded interactions involving multiple users, resembling how an individual would participate in recorded lectures and video conferences. In the original TSI framework, one of the transformations which was proposed, yet remains unstudied, is transforming time. By allowing users to join social interactions which happened in the past, we are beginning to provide a framework for understanding how time can be transformed.

2.3 Perceiving Virtual Group Social Interactions

2.3.1 Recorded vs. Generated Social Interactions. In light of the progress made in facilitating believable dialogues between virtual agents [11, 42, 51, 88, 89], another relevant body of literature focuses on generating social interactions through virtual agents. One seminal work is the social force model [35], which posits that the motion of crowds can be simulated through attractive and repelling social forces such as the force to keep a certain distance from other people and borders. Drawing on the social force model, Jan and Traum [43] presented movement simulations of conversational agents based on forces derived from the speaker, outside noise and proxemics. Bönsch et al. [12] later extended these works by proposing a classification scheme that inferred user intent based on the user's proxemics, gaze behavior, and torso orientation to decide on the future actions of the virtual agents (e.g., welcoming, rearranging, bidding farewell). When a user is determined to be joining the group of virtual agents, the agents steps back to include the user, and involves the user in their gazing strategy. Relatedly, works have also proposed methods for generating socially-acceptable and realistic trajectories for approaching conversational groups of agents [90] and examined how individuals perceived and responded to invitations for joining a group of virtual agents [40].

While our work bears similarities with works on virtual agents, there are several differences between recorded social interactions and social interactions automatically generated and scripted for virtual agents. To start, because our transformations are applied and rooted in recorded interactions, our transformations can be systematically applied to past social interactions that may be difficult to generate or are unique to specific use-cases and scenarios (e.g., class discussions, creative activities). For the same reason, applying TSI to past interactions allows users to socially connect and interact with real individuals they may be acquainted with, such as their classmates and instructors in virtual courses. This ability to experience past social interactions, which enables new students to join missed discussions and remote employees to attend past meetings in different time zones, is distinctly different from the objectives of existing works on virtual agents. Finally, in contrast to works where the dialogues between virtual agents are automatically generated, our approach preserves the verbal behavior by transforming the nonverbal behaviors of recorded interactions.

2.3.2 Real-time vs. Asynchronous Social Interactions. Finally, many compared real-time social interactions with those that are asynchronous and recorded. In particular, Kogan and Wallach [45] compared participation in group discussions to listening to taped recordings. In a study where students either interacted directly with other group members or listened to the taped audio of the discussions, the authors found that participants in the interacting groups manifested a greater extent of risky shifts following the discussion. More recently, researchers have compared real-time video content to recorded ones for training and education purposes [17, 27, 41, 49, 50, 63]. While some studies found that students preferred face-to-face interactions for real-time engagement [63] and that the use of pre-recorded lectures yielded lower exam scores for students with lower GPAs [50], many highlighted the benefits for asynchronous interactions. Specifically, researchers found that students preferred asynchronous lectures over live lectures given their convenience, educational effectiveness, and flexibility [41], and reported that students utilized recordings to supplement learning and make up for missed lectures [49]. Similarly, works have also noted that students reviewed, paused, and rewound recorded content for clarifications [63] and found themselves more likely to look up information, stay focused, and retain information in asynchronous settings [17]. For technical systems such as HyperMeeting [27] and the Collaborative Recorded Meeting [59], researchers also highlighted the need to address disorientation from navigating across multiple videos and the usefulness of shared annotations for non-attendees in information retrieval.

Given the many benefits associated with asynchronous interactions posed by prior works, we are interested in extending the discussion into immersive settings. While our work shares a similar motivation to works that envision the use of asynchronous VR interactions in future offices [23] and reliving experiences with other co-located users [83], we focus on investigating the ways we can acclimate new users into past group social interactions in VR and poses the question of how nonverbal transformations on recorded social interactions can alter perception and user behavior.

3 METHODS

3.1 Recordings of Group Discussions

To study the effects of the nonverbal behavior transformations, we recorded a group of four research assistants discussing two hypothetical scenarios selected from the Choice Dilemma Questionnaire (CDQ) [44]. Proposed by Kogan and Wallach, the CDQ consists of 12 hypothetical scenarios, each describing a character facing a choice between a riskier choice that could yield higher rewards and a more cautious choice with a higher guarantee of a moderate reward. By facilitating group discussions based on these hypothetical scenarios and measuring participant opinions before and after the discussion, researchers were able to investigate the implications of different interventions, for example choice shifts after recorded and real-time discussions [45] and the effects of persuasive arguments and social decision schemes on group decisions [94]. Scenarios D and H are chosen for the study as they yielded relatively neutral initial choices and do not require additional adaptation to the present-day context [18]. Scenario D describes a dilemma of a college football team coach choosing between a risky and conservative gameplay. In scenario H, a college senior debates between the risky path of a concert pianist or the cautious choice of becoming a physician. To account for users preferring one option over another, we recorded two discussions for each scenario, one where the group steered the conversation towards the risky option, and the other towards the cautious option.

Then, for each recording, we generated 12 arguments associated with the scenario and the option (i.e., pro-risk or pro-caution), assigned them evenly amongst three of the four research assistants, and further ordered them in a logically coherent sequence while also ensuring that each person presented one argument every three turns. The fourth research assistant acted as the moderator by opening and closing the discussion. We recorded the group discussion using ENGAGE, a social VR platform. Each discussion took around 4 minutes and 30 seconds. Additional information regarding the recording process can be found in the supplementary material.

3.2 Transforming Nonverbal Behavior in Recorded Group Discussions

3.2.1 Proxemics Manipulation. Our work leverages the preliminary design space outlined by Wang et al. [84]. Specifically, we adopted their procedure for manipulating proxemics such that the new user will be spatially accommodated into the recorded scene. Since the authors' approach places all users (the new user included) into positions evenly-spaced around a circle, this approach is appropriate for adapting our recordings of group discussion as the avatars naturally engaged in the conversation in a circle. When selecting the location of the new user, we preserved the spatial positioning of the recorded avatars relative to the new user. As seen in Figure 2, this means that the leftmost avatar in the original recording will remain the farthest left regardless of the applied proxemics transformation. In practice, this could reduce the potential confound that different relative avatar positioning can have on user perception and behavior across different proxemics manipulations. Finally, as our choice of d_{social} – the parameter that corresponds to the diameter of the circle after proxemics transformation – is condition-dependent, we detail our rationale for determining d_{social} in Section 3.2.3.

3.2.2 Gaze Manipulation. Again, referencing Wang et al. [84], we applied their procedure for remapping gaze to correct for distorted relative gaze directions following proxemics manipulations. This procedure retained the nonverbal behavior of eye contact by ensuring that a recorded avatar looking at another avatar will remain looking at that avatar despite changes made to the positions of the recorded avatars. Further, we adopted the authors' approach for inserting gaze towards the new user by sampling similar existing gaze instances. Although our recordings of group discussion are shorter in duration and thus contained a smaller variety of gaze instances to sample from, we found empirically that the method for inserting gaze based on a reference gaze instance that is similar in interpersonal distance (i.e., distance between the positions of the asynchronous user and the recorded avatar) and conversation roles (i.e., non-speaking avatar looking at a non-speaking user, speaking avatar looking at a non-speaking user) still generated believable social interactions. Finally, the authors defined $r_{transformed}$, a parameter representing the average ratio of manipulated frames to total number of frames for adding the direct gazes across all recorded avatars. Since frames in ENGAGE correspond to timestamps that are largely evenly-spaced across the recording, this parameter quantifies the percentage of total discussion time that is manipulated by a given transformation. To determine $r_{transformed}$, we piloted different values, with details provided in Section 3.2.3.

3.2.3 Conditions. As the experience of exploring (e.g., physical walking, virtual smooth translating, and teleporting around a virtual room during social interactions) differs from the act of observing from a static location, we fixed the position of the new user across all stimuli by instructing them to stand over a green circle placed at a selected location. We transformed the nonverbal behavior and created four conditions based on each recording. The four conditions differed in the extent to which we transformed the recorded social interactions, and included a baseline condition without any nonverbal transformation. Figures 1 (b–d) show top-down schematics of the three conditions with nonverbal transformations and Figure 2 shows example screenshots in ENGAGE for each of the four conditions.

Unmodified group discussion. In the unmodified condition, the new user observed the recordings as a bystander standing away from the group. To determine where users prefer to stand when observing the unmodified recordings, we conducted a pilot study (N=5). Each participant was shown one group discussion recording used for pilot tests and instructed to observe and engage in the discussion in any way they feel comfortable. In reviewing the tracking data of the participants, we noticed that while some participants teleported or smoothly translated around the virtual environment at the start of the discussion, all participants eventually picked a fixed spot away from the group to observe the remainder of the social interaction.

We thus determined the location of the new user (i.e., green circle) by first picking two avatars in the group discussion, and calculating the position for the new user that would form an equilateral triangle between that position and the root positions of the two avatars. There are two locations that satisfy this requirement, one inside and one outside the region enclosed by the avatars, and we chose the location farther from the group to align with the participant behavior observed in the pilot study. This procedure allowed us to systematically determine the asynchronous user location that conformed to the interpersonal distance within the discussion group.

Spatially accommodated group discussion. To spatially accommodate the new user, we followed the protocol introduced

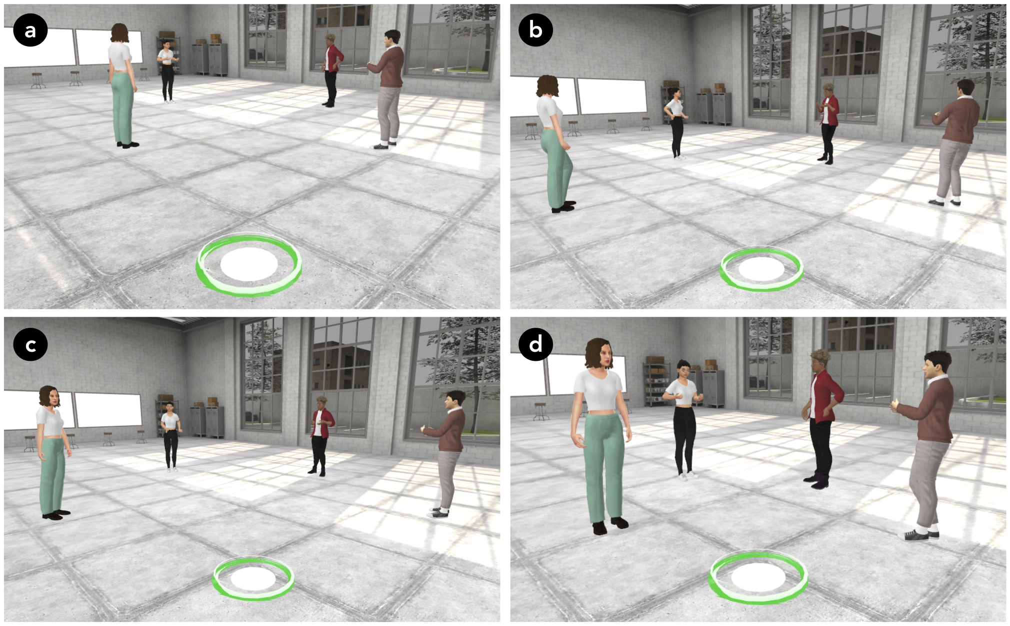

Figure 2: Example screenshots of conditions in ENGAGE where the green circles denote the position of an asynchronous user. (a) Unmodified group discussion. (b) Spatially accommodated group discussion. (c) Spatially accommodated group discussion with added gaze. (d) Close-range and spatially accommodated group discussion with added gaze.

in Section 3.2.1 and determined the diameter of the circle by first calculating the interpersonal distance of the original recording. Interpersonal distance was measured by taking the mean of the distances from each of the four recorded avatars to the recorded avatar standing closest to them (M=4.03, SD=0.52). Using this value, we calculated the value of the diameter that would yield the same interpersonal distance as the original group discussion, now with the addition of the new user. We then transformed the proxemics of the recording using the derived diameter and applied gaze remapping scheme to correct for distortions of gaze directions after proxemics transformations. Effectively, this condition opened up the space to accommodate the new user into the circle, whilst retaining the same interpersonal distance as in the original recording.

Spatially accommodated group discussion with added gaze. This condition builds on the previous condition by taking the transformed social interactions and further adding gaze towards the new user using the procedure proposed in Section 3.2.2. To increase the realism of added gaze towards the new user, we set the ceiling of the created gaze dataset to five seconds, with the floor set to one second. Through piloting various $r_{transformed}$ values, we decided to use a ratio of 0.4 so that the added gazes are noticeable but not overwhelmingly frequent for a new user who does not actively contribute to the discussion.

Close-range and spatially accommodated group discussion with added gaze. The final condition is similar to the previous condition, with the exception of choosing a diameter value that yielded an interpersonal distance smaller than that of the original. Specifically, we chose the diameter value that reduced the original interpersonal distance by 50 percent (M=2.01, SD=0.26). We chose this ratio so that the interpersonal distance remained within the social zone based on Hall's proxemics theory [31]. While the original interpersonal distance falls on the higher end of the social distance range of 48 inches to 12 feet, this condition tightens the interpersonal distance to the lower more intimate end of the spectrum. Similar to the previous condition, we also added in gaze towards the new user following Section 3.2.2 with the $r_{transformed}$ set to 0.4.

4 USER STUDY

We conducted a between-participant user study comparing the four conditions, namely users watching group recordings that were (1) *unmodified*, (2) *spatially accommodated*, (3) *spatially accommodated with added gaze*, and (4) *close-range and spatially accommodated with added gaze*. Specifically, each participant is shown two recorded group discussions in the same condition, one for each of the two chosen choice dilemma scenarios. Whether each of the discussion is pro-caution or pro-risk and which scenario is shown first to each participant are randomly assigned using a Latin Square Design, where the order and main argument for each discussion are counterbalanced. We present specific hypotheses and research questions in Section 4.4.

4.1 Participants

A power analysis of our regression model as described in Section 5, including $\alpha = 0.05$, power = 0.95, and $f^2 = 0.15$ for detecting medium effect size [20], showed a sample size of 138. We fell short of this recruitment goal due to logistics and planning constraints. With this in mind, we recruited 128 participants (72 female, 55 male, 1 preferred not to answer) between the age of 18 and 68 (M=24.26,

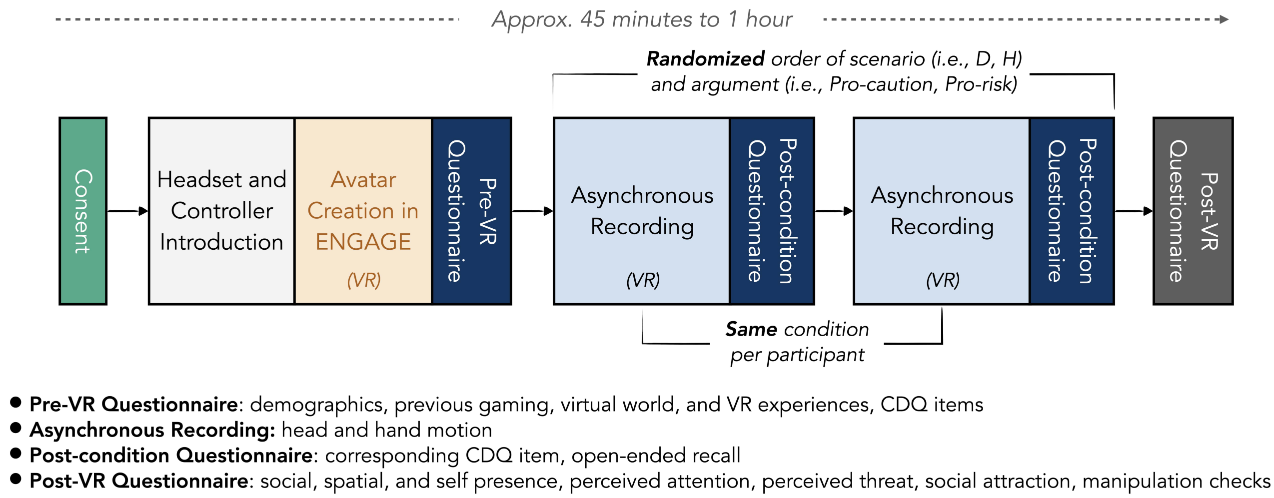

Figure 3: Overview of user study procedure. At the bottom of the figure, bullet points detail the measurements taken at each stage of the user study. We omit bullet points for stages of the user study where no measurement is taken.

SD=6.30) from a medium-sized private university through the department participant pool, email, and word-of-mouth. Participants were allowed to select more than one category for race and ethnicity. 7 selected African American, 72 selected Asian, 32 selected Caucasian, 16 selected Hispanic, and 6 selected other. The majority of the participants (96 out of 128) had used VR before. Regarding virtual worlds and social virtual environments, 73 reported never visiting, 40 visiting rarely, 13 visiting sometimes, and 2 visiting several times a week. In terms of gaming experience, 35 reported never playing, 48 playing rarely, 33 playing sometimes, and 12 playing several times a week or daily. The study procedure was approved by the university IRB board, and consent was obtained before each participant began the study.

4.2 Procedure

Figure 3 shows an overview of our study procedure. After obtaining consent from each of the participants, the research personnel introduced each participant to the hand-held controllers and the Meta Quest 2 headset. The participant then put on the headset and was instructed to create an avatar that looks and feels like them using the ENGAGE platform. Following this, each participant was led to another room, where they completed a questionnaire on a desktop computer on demographics, previous gaming, virtual world and VR experiences, as well as their initial selections on the two choice dilemma scenarios (i.e., D and H). Each participant was then guided back to the main laboratory space and shown the first recorded discussion in headset. Following the first recording, they were guided back to the survey room to complete a questionnaire asking them their post-VR selection on the choice dilemma scenario the recording discussed as well as open-ended recall questions on what each of the avatars said during the discussion. Upon completing the questionnaire, the participant viewed the second recording in the main laboratory space, after which they completed another questionnaire formatted the same way as the one before. Each participant then completed one final questionnaire that included questions on their VR experience, followed by questions on manipulation checks.

4.3 Measures

We collected subjective measurements through questionnaires on social, spatial, and self presence, perceived attention, perceived threat, and social attraction. We referenced Han et al. [32] for the presence questionnaire items and Harms and Biocca [33] and Fauville et al. [22] for the items on perceived attention, perceived threat, and social attraction. We recorded the items on a 5-point likert scale from “Not at all” to “Extremely”. A composite was created for each construct measured after ensuring a sufficiently high reliability using the Cronbach's alpha test or Spearman-Brown formula, by extracting the mean across the related items. Since the Spearman-Brown reliability for perceived threat is < 0.7, we analyzed the two items separately. In comparison to our manipulation check questions on gaze that asked participants to estimate the amount of direct gaze they received from each avatar, the items on perceived attention are different as being looked at by avatars may not directly correlate with a participant's perceived attention from and towards the avatars. Table 1 shows all subjective questionnaire items as well as their reliability measurements. Additionally, we collected responses to open-ended recall questions on what each avatar said in a given scenario.

We also recorded the tracking data of participants, specifically the position (x, y, z) and orientation (roll, pitch, yaw) of the headset and the two controllers over time. The user motion data was recorded at roughly 30 Hz. Taken together with the tracking data from the discussions that recorded both nonverbal and verbal behavior, the tracking data from participants allowed us to derive additional behavioral measurements. For the remainder of this section, we describe each of the behavioral metrics and provide justifications for including them in the exploratory research questions discussed next.

Table 1: Post-VR subjective questionnaire items. We report reliability measures for constructs with more than two items using Cronbach's α and perceived threat measurement using Spearman-Brown.

Construct	Reliability	No.	Question
Social presence	$r_\alpha = 0.85$	1.1	It felt like the other people in the room were with me.
		1.2	It felt like I was face-to-face with others.
		1.3	It felt like the other people were aware of my presence.
Spatial presence	$r_\alpha = 0.76$	2.1	It felt as if I was inside the virtual world.
		2.2	It felt as if I was visiting another place.
		2.3	It felt like I could reach out and touch the objects in the virtual environments.
Self presence	$r_\alpha = 0.72$	3.1	I felt that my avatar represented me.
		3.2	When something happened to my avatar, I felt like it was happening to me.
		3.3	I felt like I was able to control my avatar as though it were my own.
Perceived attention	$r_\alpha = 0.71$	4.1	I felt like the group paid close attention to me.
		4.2	I paid close attention to the group.
		4.3	I think the group remained focused on me throughout our interaction.
		4.4	I remained focused on the group throughout our interaction.
Perceived threat	$r_{SB} = 0.35$	5.1	I think the group is dominant.
		5.2	I think the group is threatening.
Social attraction	$r_\alpha = 0.86$	6.1	If I had a video conference with the group, I would like this group.
		6.2	If I had a video conference with the group, I would get along with this group.
		6.3	If I had a video conference with the group, I would enjoy a casual conversation with this group again.
		6.4	If I had a video conference with the group, I would think this group is friendly.

4.3.1 Interpersonal Distance. As interpersonal distance is closely related to one's perceived level of intimacy [3], we are interested in understanding how interpersonal distance is related to subjective measurements such as perceived threat and social attraction. We measured interpersonal distance by calculating the average distance between a participant's head position and the head positions of all four recorded avatars across the duration of the entire discussion. Unlike the manipulation check question that asks the participants to estimate their distance to the closest avatar, this metric takes into account the relative spatial positioning of all recorded avatars. While we explicitly instructed the participants to stand over a pre-selected location, average interpersonal distance factors in more subtle movements such as head translation and minor paces that participants can take near the pre-selected location.

4.3.2 Direct and Mutual Gaze. We calculated two metrics related to participants' gaze patterns, namely direct gaze and mutual gaze. Both metrics are important to track as past work have highlighted eye contact (i.e., mutual gaze) as an important predictor of interpersonal distance under the equilibrium theory [3, 8], while the unidirectional action of looking towards others (i.e., direct gaze) can provide insights for visual and conversational attention [56, 82].

To derive a measurement for direct gaze, we calculated the percentage of time each participant spent looking at the recorded avatars. Following the thresholding technique proposed by Miller et al. [57], we defined direct gaze as a participant's head orientation falling within both 15 degrees of yaw and pitch angle of any of the four recorded avatars. We chose the threshold value of 15 as the angle between one's gaze and head direction is typically less than 15 degrees [24]. In addition to reporting results using the 15 degree threshold, we also analyzed gaze behavior using lower threshold values and contextualized the differences and similarities when interpreting the results in Section 6.1.2. Details on the threshold value comparisons can be found in the supplementary material.

We sampled the tracking data at 10 Hz. We chose this sampling rate since we manually aligned the tracking data of the recorded discussions to that of the participants using audio recordings, with an error margin of roughly 0.1 seconds. This method of measuring direct gaze introduced one nuance, as the participants in the *unmodified* condition were not standing in a circle with the rest of the recorded avatars, but rather farther away from the group as seen in Figure 2 (a). With these participants being able to fit more avatars into their field of view, we interpret the results related to the *unmodified* condition with caution. Furthermore, as multiple avatars can be within 15 degrees of yaw and pitch of a participant's head orientation, especially in the *unmodified* condition, the summation of ratios for gaze towards avatars can exceed 100 percent. Relatedly, we quantified mutual gaze as the ratio of time each participant spent gazing at an avatar who is also looking at them. Therefore, we defined mutual gaze as a participant's head orientation falling within 15 degrees of yaw and pitch of an avatar, and that avatar's head orientation also falling within 15 degrees of yaw and pitch of the participant.

4.3.3 Speaking Role. Speaking role refers to whether a recorded avatar is speaking at a particular timestamp. This is an important factor to consider as gaze can carry different functionalities depending on the conversation roles (e.g., speaker, listener) of the individuals [4]. Past works have shown that turn-taking is closely related to eye gaze patterns and conversational attention [73, 82] and that characteristics such as loudness of speech [79] can further influence interpersonal distance as users seek to maintain an appropriate level of intimacy. For these reasons, we further separate metrics related to gaze patterns into direct gaze and mutual gaze towards non-speaking and speaking avatars, yielding a total of four different gaze-related measurements.

4.3.4 Head Movement. Finally, we break down head motion data into head rotation (i.e., pitch, yaw, roll) and head translation (i.e., movement in 3D space). The choice of analyzing head rotation in the three axes separately is motivated by past works suggesting that head rotation on different axes corresponds to different cognitive processes. For example, works have shown that yaw is closely related to anxiety towards virtual agents [87], presence [37, 74], and mental and cognitive demand [37], while pitch was effective in predicting risk perception [72]. In this work, we defined head rotation as the average change in degrees per second for a given axis, whereas the amount of head translation is measured in change in head position in meters per minute. We sampled the tracking data at 10 Hz when calculating both metrics.

To facilitate our understanding of whether conditions affect direction of head motion, we complemented our single-value head translation measurement by further breaking down head translation into two orthogonal components relative to the participant's current head orientation, one in the forward-backward direction and another in the sideways direction. Specifically, at each timestamp for a given recording, we calculated the magnitude of the projection of the participant's head motion vector in the horizontal plane onto the head's forward vector (i.e., the unit vector the head is facing), as well as the magnitude of its residual. The sum of the magnitude for the projection, averaged over total time of a recording, quantifies the amount of forward-backward head motion, whereas that for the residual quantifies the amount of sideways head motion.

4.4 Hypotheses and Research Questions

We pre-registered hypotheses using OSF[1]. In the result section, we present findings to three hypotheses on self-reported social presence **(H1)**, perceived attention **(H2)**, and direct gaze **(H3)**[2]. Specifically, we hypothesize that compared to the *unmodified* condition, each of the three transformed conditions will yield higher values of social presence **(H1a)**, perceived attention **(H2a)**, and a greater amount of direct gaze towards both speaking and non-speaking avatars **(H3a)**. However, we do not expect the *spatially accommodated* condition to be significantly different from the *unmodified* condition in terms of social presence **(H1b)**, perceived attention **(H2b)**, and direct gaze **(H3b)**. Finally, we hypothesize that compared to participants in the *unmodified*, *spatially accommodated*, and *spatially accommodated with added gaze* conditions, participants in the *close-range and spatially accommodated with added gaze* condition will report significantly higher social presence **(H1c)**, greater perceived attention **(H2c)**, and higher amount of direct gaze **(H3c)**.

We further formulated three research questions based on the behavioral measurements. They were:

- **RQ1 - Interpersonal Distance**: how is average interpersonal distance related to subjective measurements associated with an individual's perceived level of intimacy, namely perceived threat and social attraction?
- **RQ2 - Gaze**: how do different transformations of asynchronous social interactions influence participant mutual gaze patterns, and how do gaze patterns differ between direct and mutual gaze, speaking roles, and relative positioning of the recorded avatars?
- **RQ3 - Head Movement**: how do different transformations of asynchronous social interactions influence participant head motion such as the amount of rotation and translation?

[1]https://osf.io/d2pzt/

[2]The supplementary material contains results on the pre-registered hypothesis on persuasion.

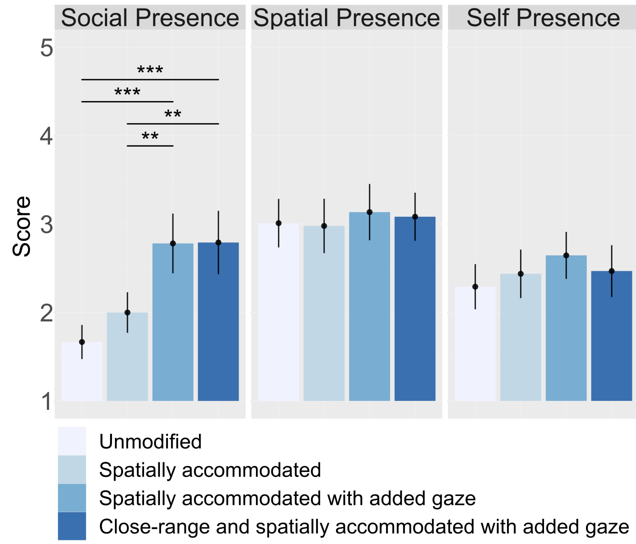

Figure 4: Presence scores. Whiskers show confidence intervals and significant post-hoc comparisons are shown (p > .05, *: $p \le .05$, **: $p \le .01$, ***: $p \le .001$). Conditions with added gaze yielded higher social presence scores compared to those without added gaze. There was no significant difference across conditions for spatial presence and self presence.

Are the patterns in head translation consistent if we further break down the direction of head translation into the forward-backward and sideways motion components?

5 RESULTS

We conducted manipulation checks for gaze and proxemics. Using a Welch's t-test, we confirmed that those in the conditions with added gaze felt more looked at by the recorded avatars (M=25.11, SD=14.54) than those in conditions without added gaze (M=8.68, SD=9.64), $t(109.42)=7.53$, $p<.001$. Similarly, a Welch's t-test also confirmed that participants in the *close-range and spatially accommodated with added gaze* condition perceived the closest avatar to be standing closer to them (M=3.28, SD=1.33) than those in the other three conditions (M=4.91, SD=1.89), $t(75.73)=-5.36$, $p<.001$.

We next report key findings related to our hypotheses as well as exploratory analysis on our research questions. To analyze the effect of conditions on the dependent variables, we built linear models predicting each of the dependent variables using the condition and participant-specific predictors (i.e., age, ethnicity, gender, and VR experience). For dependent variables that were derived over tracking data and were thus also stimuli-specific, we added the same set of participant-specific predictors and also stimuli-specific predictors, namely scenario (i.e., D or H), argument (i.e., pro-caution or pro-risk), and order (i.e., shown first or second). We evaluated statistical significance at $\alpha = 0.05$. In cases when there is a main effect of condition, we conducted additional post-hoc analyses with Bonferroni correction. Finally, we report the main effects of the participant-specific predictors (i.e., demographics). For demographic variables that are either continuous or binary (i.e., age, gender, VR experience), we discuss the direction of the

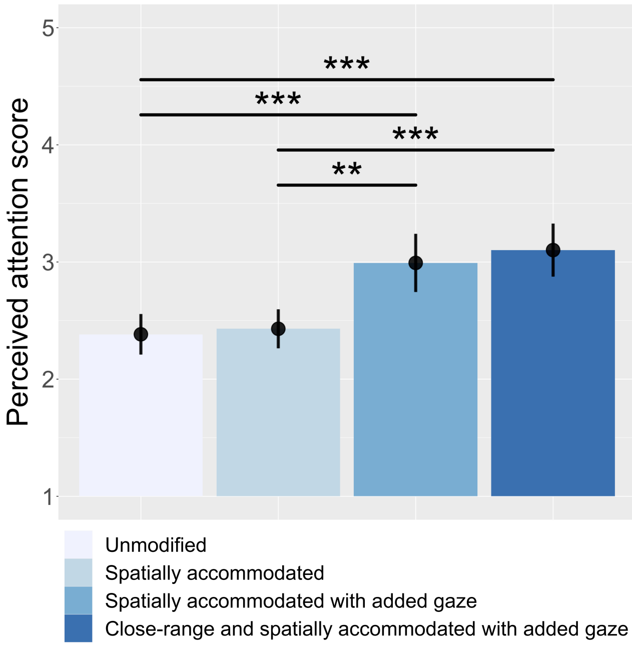

Figure 5: Perceived attention. Whiskers show confidence intervals and significant post-hoc comparisons are shown (p > .05, *: $p \leq .05$, **: $p \leq .01$, ***: $p \leq .001$). Conditions with added gaze yielded higher perceived attention scores compared to those without added gaze.

main effects. Since the patterns for ethnicity were fairly complex and not the focus of the work, we highlight the main effects of ethnicity when presenting results and expand on these findings in the supplementary material.

5.1 Subjective Measurements

5.1.1 Presence. Figure 4 shows our results on presence. For social presence, there was a significant difference across conditions (p<.001) and VR experience (p=.01). Specifically, we found that participants with VR experience reported lower social presence compared to those without VR experience. A post-hoc analysis revealed that the *unmodified* condition yielded lower social presence scores than the *spatially accommodated with added gaze* and the *close-range and spatially accommodated with added gaze* conditions (ps<.001). Similarly, we also found that the *spatially accommodated* condition also yielded lower social presence scores than the *spatially accommodated with added gaze* condition (p=.005) and the *close-range and spatially accommodated with added gaze* condition (p=.002). Compared to the *unmodified* condition, the *spatially accommodated* condition increases social presence, though not significantly (p=.61). **H1** is thus partially supported. We found no significant differences for spatial presence and self presence across the four conditions (p=.87 for spatial presence, p=.34 for self presence). For both measures, there were no significant participant-specific predictors.

5.1.2 Perceived Attention. As shown in Figure 5, there was a significant difference across conditions for perceived attention (p<.001). There was no significant participant-specific predictor. Our post-hoc analysis showed that those in the *unmodified* condition reported lower perceived attention than those in the *spatially accommodated with added gaze* condition (p<.001) and *close-range and spatially accommodated with added gaze* condition (p<.001). We also found that participants in the *spatially accommodated* condition reported lower perceived attention than those in the *spatially accommodated with added gaze* condition (p=.008) and *close-range and spatially accommodated with added gaze* condition (p=.001). **H2** is thus partially supported.

5.1.3 Perceived Threat. The Spearman-Brown reliability for the two questionnaire items on perceived threat is 0.35. We thus reported the results for the two items separately. For item 1, we found that compared to female participants, male participants reported lower perceived threat (p=.01). For both items on perceived threat, we found no significant difference across conditions, though the *unmodified* condition yielded the lowest perceived threat scores on average (M=1.94 for item 1, M=1.06 for item 2) and the *close-range and spatially accommodated with added gaze* condition the highest (M=2.53 for item 1, M=1.34 for item 2). This suggested an inverse relationship between interpersonal distance and perceived threat (**RQ1**). To confirm this, we created two additional linear models, each predicting one of the two perceived threat measurements using the average interpersonal distance, condition, as well as the participant-specific and stimuli-specific predictors. Both models showed a main effect of average interpersonal distance (p=.02 for item 1 and p=.002 for item 2), with closer average distance predicted to have higher perceived threat.

5.1.4 Social Attraction. We found no significant difference in social attraction scores across the four conditions (p=.93). There was a main effect of VR experience (p=.04), where participants with VR experience reported lower social attraction than those without. To investigate the relationship between social attraction and interpersonal distance (**RQ1**), we again created a linear model predicting social attraction using the average interpersonal distance, condition, as well as the participant-specific and stimuli-specific predictors. We found no main effect of average interpersonal distance in predicting social attraction (p=.40).

5.1.5 Recall. We coded the open-ended content recall question responses using a coding scheme derived from the four sets of arguments used to generate each of the recordings. Specifically, for each of the four recordings, we derived a coding rubric with 13 points, 1 point for each of the 12 arguments, and another for mentioning that the correct avatar was moderating the conversation. The items are each graded in a binary basis and the content recall score is calculated as the sum of the scores attributed to the 13 items. The notion of measuring recall as the number of correctly recalled elements is similar to those found in [5, 39, 47, 60]. Similar to Baceviciute et al. [5], two coders graded all of the responses independently before meeting to discuss each of the disagreements in the content recall score, updating their scores if deemed appropriate. After calibrating their scores, the coders reached a full consensus (i.e., the same content recall score) on 234 out of the 256 coded responses. We then averaged the two scores as the final content recall score. We calculated Intraclass correlation (ICC) to evaluate the reliability

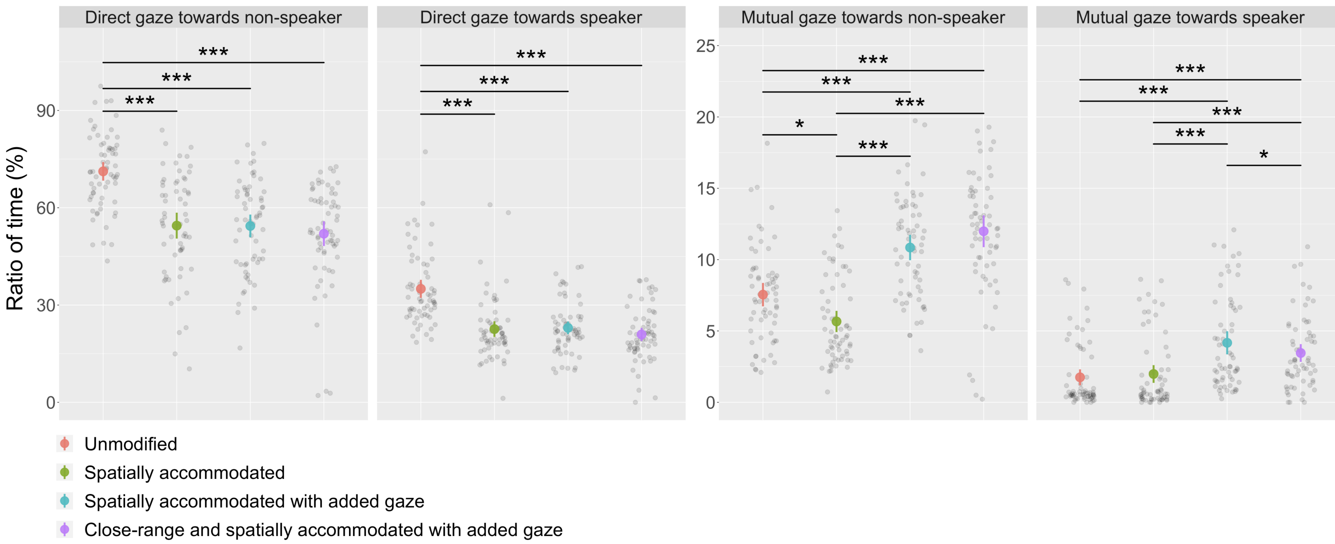

Figure 6: Gaze breakdown by type and speaking role. The left and right panels show direct gaze and mutual gaze, respectively. Whiskers show confidence intervals and significant post-hoc comparisons are shown (p > .05, *: $p \le .05$, **: $p \le .01$, ***: $p \le .001$). The *unmodified* condition yielded higher amounts of direct gaze compared to the three transformed conditions. Conditions with added gaze yielded high amounts of mutual gaze compared to conditions without added gaze.

Table 2: Intraclass correlations for each of the four content recall coding schemes before and after calibration.

Stimuli	Pre-calibration ICC	Post-calibration ICC
Scenario D, Pro-caution	0.76	0.96
Scenario D, Pro-risk	0.73	0.99
Scenario H, Pro-caution	0.84	0.99
Scenario H, Pro-risk	0.91	0.99

of the coding scheme by using a Two-Way Mixed Effects Model, evaluating for consistency. Table 2 shows the ICC scores for the four stimuli across before and after calibration.

Using a linear model predicting the final content recall score using condition, participant- and stimuli- specific predictors, we found no significant difference across the four conditions (p=.76). There were main effects for scenario (p<.001), gender (p<.001), and VR experience (p=.05). Specifically, we found that participants remembered more when watching the discussions for scenario H when compared to scenario D and that female participants remembered more arguments than male participants. Compared to participants without VR experience, those with VR experience remembered less arguments.

5.2 Behavioral Measurements

5.2.1 Direct Gaze. We analyzed how the four different conditions influence direct gaze patterns by speaking roles of the recorded discussion by building models predicting direct gaze with condition, participant- and stimuli- specific predictors. There were significant differences across conditions for direct gaze towards both non-speakers and speakers (ps<.001). As shown in the left two panels of Figure 6, we found that in comparison to participants in the three transformed conditions, those in the *unmodified* condition spent a significantly greater ratio of time looking at both speakers and non-speakers. Our post-hoc analysis showed no other significant difference across the three transformed conditions. **H3** is not supported. For the model predicting direct gaze towards non-speaker, we found main effects of age (p=.04) and ethnicity (p=.02), where older participants spent more time looking at non-speakers.

5.2.2 Mutual Gaze. To investigate **RQ2**, we analyzed whether the condition has an impact on the percentage of mutual gaze towards speaking and non-speaking avatars. The right two panels of Figure 6 summarize our findings. In both linear models predicting mutual gaze towards non-speaking and speaking avatars respectively, there was a main effect of condition (p<.001). For the model predicting mutual gaze towards speaking avatars, we also found a main effect of ethnicity (p<.001).

In the post-hoc analysis, we found that the mutual gaze towards non-speaking avatars for the *unmodified* condition was higher than the *spatially accommodated* condition (p=.03), and lower than the *spatially accommodated with added gaze* (p<.001) and the *close-range and spatially accommodated with added gaze* conditions (p<.001). Mutual gaze towards non-speaking avatars for the *spatially accommodated* condition was also lower than that for the *spatially accommodated with added gaze* (p<.001) and *close-range and spatially accommodated with added gaze* conditions (p<.001).

The post-hoc analysis also showed that participants in the *spatially accommodated with added gaze* condition spent a higher ratio of time engaging in mutual gaze with speaking avatars than the *unmodified* (p<.001), *spatially accommodated* (p<.001), and *close-range and spatially accommodated* conditions (p=.04). Additionally, participants in the *close-range and spatially accommodated* condition also spent a greater ratio of time engaging in mutual gaze with

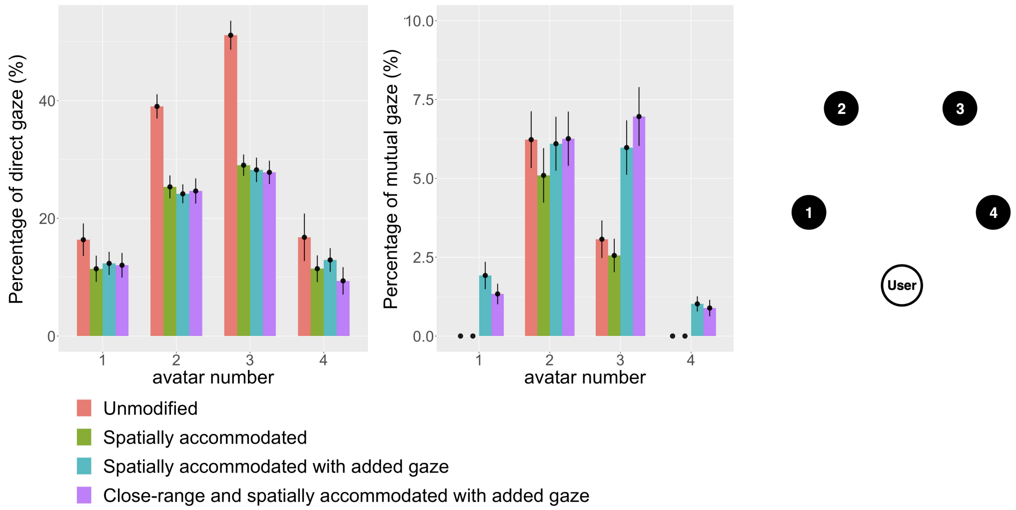

Figure 7: Gaze breakdown by avatar. Left and middle panels show the percentage of direct gaze and mutual gaze, respectively. Right panel shows the schematic of where each of the four avatars are positioned in relation to the new user. Whiskers show confidence intervals. Participants engaged in more direct and mutual gaze with avatars 2 and 3 than avatars 1 and 4. Participants in conditions without added gaze did not engage in mutual gaze with avatars 1 and 4.

speaking avatars when compared to the *unmodified* (ps<.001) and *spatially accommodated* conditions (ps<.001).

5.2.3 *Gaze by Avatars.* We were interested in understanding how the positioning of the avatars within an asynchronous conversation affects gaze (**RQ2**). To do this, we took the direct and mutual gaze measurements calculated from the previous sections and further plotted out the spread of gaze across the four recorded avatars, where avatar 2 is the moderator of the discussion (Figure 7). We built two additional linear models predicting the amount of direct gaze and mutual gaze using the avatar number, as well as the participant-specific and stimuli-specfic independent variables as predictors. In both cases, there were significant differences across the four avatar numbers (ps<.001) and ethnicity (ps=.05). For both models, the post-hoc analysis revealed significant differences across all pairwise comparisons of the avatar number (ps<.001), with the exception for the comparisons between avatar 1 and avatar 4 (p=1 for direct gaze and p=.74 for mutual gaze).

5.2.4 *Head Rotation.* To gain a better understanding of whether condition has an effect on the amount of head rotation (**RQ3**), we built a linear model predicting head rotation for each of the three axes (i.e., pitch, yaw, roll). For all three axes, we found main effects of ethnicity (ps<.001) and gender (ps<.001), where male participants exhibited more head rotation compared to female participants. There was a main effect of VR experience when predicting pitch and yaw head rotations (ps=.04), where participants with VR experience exhibited more head rotation compared to those without.

Figure 8 summarizes our findings regarding head rotation across conditions. There were significant differences across the four conditions for roll (p=.002) and yaw (p<.001), but not for pitch (p=.61). For yaw, the results showed that the rotation speed for sessions in the *unmodified* condition was slower than each of the transformed

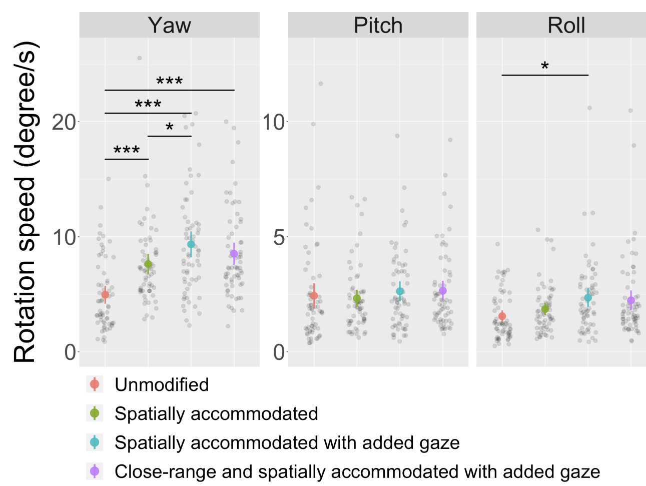

Figure 8: Head rotation. Whiskers show confidence intervals and significant post-hoc comparisons are shown (p > 0.05, *: $p \leq 0.05$, **: $p \leq 0.01$, ***: $p \leq 0.001$). Participants in the *unmodified* condition engaged in slower head yaw rotation compared to the three transformed conditions. Participants in the *spatially accommodated with added gaze* condition also exhibited faster head yaw rotation compared to the *spatially accommodated* condition, and head roll rotation compared to the *unmodified* condition.

three conditions (ps<.001). This is reasonable as the avatars themselves take up a smaller field of view in the *unmodified* condition and thus requires less head yaw rotation to look at different avatars. Further, sessions conducted in the *spatially accommodated with added gaze* condition had faster yaw rotation speed than those in

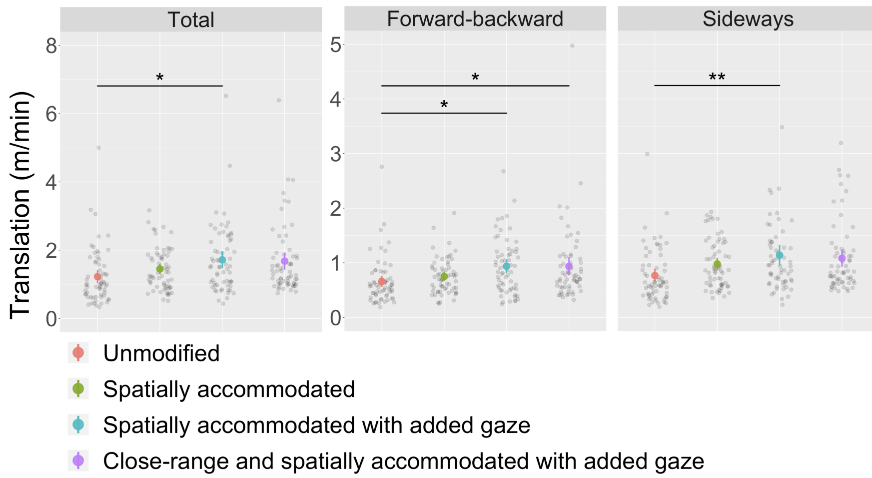

Figure 9: Head translation. Whiskers show confidence interval and significant post-hoc comparisons are shown (p > 0.05, *: $p \leq$ 0.05, **: $p \leq$ 0.01, ***: $p \leq$ 0.001). Compared to participants in the *unmodified* condition, those in the *spatially accommodated with added gaze* condition exhibited greater head translation in terms of total, forward-backward, and sideways motion. Participants in the *close-range and spatially accommodated with added gaze* condition exhibited greater forward-backward head translation compared to those in the *unmodified* condition.

the *spatially accommodated* condition (p=.05). For roll, we found that the rotation speed for sessions in the *spatially accommodated with added gaze* condition was significantly faster than those in the *unmodified* condition (p=.01).

5.2.5 Head Translation. Finally, we analyzed whether the condition has an impact on the amount of head translation for participants (**RQ3**). In our linear model predicting head translation, we found that there was a significant difference across conditions (p<.001). As seen in the left panel of Figure 9, our post-hoc analysis revealed that the amount of head translation for user study sessions in the *unmodified* condition was smaller than those in the *spatially accommodated with added gaze* condition (p=.01).

To further understand how directions of head motion contributed to this difference, we conducted an additional analysis breaking down head motion into two orthogonal components, one in the forward-backward direction and another in the sideways direction. Shown in the middle and right panels of Figure 9, there were significant differences in head translation across the four conditions (ps<.001) in both directions. Sessions in the *unmodified* condition exhibited less head translation than the *spatially accommodated with added gaze* condition in both the forward-backward direction (p=.02) and sideways direction (p=.004). Additionally, sessions in the *unmodified* condition also had less head translation than the *close-range and spatially accommodated with added gaze* condition in the forward-backward direction (p=.03).

The three models predicting total, forward-backward, and sideways head translation also revealed main effects of ethnicity (ps<.001) and gender (ps<.001), where male participants exhibited more head translation compared to female participants. For total and sideways head translation, we also found a main effect of VR experience (ps=.03), where participants with VR experience exhibited more head translation compared to those without.

6 DISCUSSION

6.1 Summary of Results

6.1.1 Social Presence and Perceived Attention. The findings on social presence and perceived attention show that participants in conditions with added eye gaze (i.e., *spatially accommodated with added eye gaze, close-range and spatially accommodated with added eye gaze*) feel a greater sense of social presence and perceived attention, while merely spatially accommodating the user does not. As we did not test a condition with added eye gaze without spatially accommodating the new user, it is unclear whether the increase in social presence and perceived attention is caused by the added eye gaze alone, or due to the combination of spatial accommodation and inserted eye gaze. Although we had hypothesized that spatially accommodating the new user at a closer range within the social zone based on Hall's proxemics theory [31] would further increase social presence and perceived attention (**H1**), our findings suggest otherwise. One explanation is that the interpersonal distance amongst the recorded avatars was already within the social zone – the appropriate interpersonal distance for interacting with acquaintances and strangers – so our approach for decreasing interpersonal distance may not have had its desired effect. Nevertheless, we believe that such proxemics manipulations could be beneficial for transforming recorded social interactions when the interpersonal distance is outside the social zone. However, as shown in Section 5.1.3, there is the risk of increasing perceived threat when decreasing the interpersonal distance.

6.1.2 Behavioral Patterns of Gaze and Head Motion in Asynchronous Social VR. Contrary to **H3**, participants in the transformed conditions did not look at recorded avatars (both speaking and non-speaking avatars) differently in terms of duration. As noted in the supplementary material, this remained the case for lower eye gaze thresholds of 5 and 10 degrees. That said, unlike our results on direct gaze, the findings on mutual gaze matched up closely to the significant post-hoc contrasts we found for social presence and perceived attention (**RQ2**). The finding that conditions with added gaze exhibited greater mutual gaze were largely consistent for lower eye gaze thresholds. We interpret this as the increase in social presence and perceived attention being closely related to the increase in mutual gaze, as opposed to direct gaze. One explanation for the differences between eye gaze thresholds is that we used a 15 degree threshold for determining the set of gaze instances to sample from when inserting gaze to the original recordings. Because of this, using different and lower eye gaze thresholds to determine participant gaze behavior (i.e., 5 and 10) can underestimate the amount of gaze we intended to add towards the new user, and yield misleading downstream estimates of mutual and direct gaze.

Additionally, the finding that participants in the *spatially accommodated with added gaze* condition spends a longer time engaging in mutual gaze with speaking avatars than they do in the *close-range and spatially accommodated with added gaze* condition can be explained by the equilibrium theory [3], which states that individuals will seek to maintain an adequate level of intimacy in social interactions through behavioral changes such as gaze and proxemics. When a speaking avatar is looking directly at the new user, the new user may feel more inclined to look away when the speaking avatar is also standing at a close interpersonal distance. This finding was consistent across all tested eye gaze threshold values, though only marginally significant for the 10-degree threshold (p=.08).

In regards to **RQ2**, our results on gaze breakdown by avatars revealed that participants engaged in more direct and mutual gaze with the two farther avatars across all conditions. This was the case for all tested eye gaze thresholds. We offer two explanations for this observation. To start, according to the equilibrium theory, an individual will tend to look more towards and engage in more eye contact when they are standing farther away from an individual. The second and simpler explanation is that users in VR may not feel inclined to move their heads in VR [66], and thus resort to looking straight ahead as opposed to the two avatars on the side. The finding that participants did not engage in mutual gaze (i.e., 0 percent) with the left- and right- most avatars in conditions that did not add in gaze is also reasonable as the two avatars never turned towards the location of the new user during the original discussion and would therefore always have their backs towards the new user.

The result that participants in the *spatially accommodated with added gaze* condition had greater yaw movement than those in the *spatially accommodated* condition suggests that adding gaze drives the participants to be more active when observing the recorded conversation. We also observed a significant increase in forward-backward head motion in the *spatially accommodated with added gaze* and *close-range and spatially accommodated with added gaze* conditions when compared to the *unmodified* condition. We hypothesize that individuals who were spatially accommodated while receiving eye gaze were more inclined to engage in actions with more forward-backward motion such as nodding, leaning in towards the speaker, and moving back during mutual gaze. That said, because we found no significant pairwise comparison between the three transformed conditions for forward-backward head translation, it is unclear how interpersonal distance and spatial accommodation independently influence this behavior. Finally, the finding that participants moved their head more (i.e., translation) in the *spatially accommodated with added gaze* condition when compared to the *unmodified* condition regardless of direction suggests that the combination of adding gaze and spatial accommodation increases activeness, and that its effect on head translation is isotropic.

6.2 Implications for Transforming Asynchronous Social Interactions in VR

6.2.1 Design Implications. We outline actionable design implications for how practitioners such as educators and organizers of virtual meetings can incorporate our proposed transformations into asynchronous scenarios.

- **When transforming recordings of a VR group discussion, consider applying both proxemics and gaze transformations.** Although our results showed that spatial accommodation alone did not increase social presence and perceived attention, the combination of spatial accommodation and added gaze did. In line with past works in augmenting synchronous social interactions [68, 80], our findings showed that transformations of asynchronous VR social interactions can also be beneficial to the user's subjective experience.
- **When determining the amount of gaze or interpersonal distance for the transformation, reference the existing gaze and proxemics behavior from the original recordings.** In our study, we observed no adverse effect in maintaining the same interpersonal distance and sampling existing gaze instances. That said, as interpersonal distance may vary depending on factors such as agent appearance [38] and cultural differences [78], we believe that one should consider the relationship between the asynchronous participant and the recorded avatars, and make adjustments to the parameters if necessary. For example, when acclimating a student into a recorded lecture, one might consider maintaining a farther interpersonal distance from the lecturer, and a closer interpersonal distance to students whom the student is more familiar with.
- **When deciding the spatial juxtaposition of the new user, consider employing mechanisms for directing user attention towards the periphery or placing avatars of greater significance near the center of the user's viewing perspective of the recorded interaction.** This is because our results indicated that participants look more at the two avatars standing closer to the center of their field of view if they looked directly towards the center of the group, regardless of conditions. To do this, one can for example adapt the utterances of the conversation depending on the user's position [61], or dynamically manipulate avatar positions based on user motion [46].
- **When setting up transformations of VR social interactions, outline key objectives and supplement our transformations with techniques that can optimize additional metrics.** As we did not find significant differences in content recall across the four conditions, it is unclear whether the proposed transformations alone can be beneficial to metrics such as learning and information recall. Thus,

depending on the specific use-case, one can consider employing approaches for recording educational content [13, 58] and for facilitating learning [70, 76].

6.2.2 Ethical Implications. In addition to considering the design implications for transforming asynchronous VR social interactions, practitioners must also consider the important ethical implications when applying these transformations. Specifically, VR users may not consent to their recordings being altered, and new users joining transformed recordings may not be aware of these manipulations and therefore experience deception. Furthermore, altering the nonverbal behavior could distort the original user intentions and lead to unintended downstream effects. In our user study, we mitigated these potential ethical concerns through (1) obtaining consent for manipulating the nonverbal behavior of the four research assistants who acted as confederates in our recorded discussions, (2) choosing hypothetical discussion topics and arguments that are impersonal and less likely to be misinterpreted after nonverbal transformations, and (3) debriefing participants of the manipulations after the study.

In line with these actions, and in the context of existing tools that augment eye contact for video conferencing [48, 52], we believe that practitioners should adopt a consent model. In particular, the model should inform and seek consent from the recorded users to transform their nonverbal behavior and provide the recorded users adequate control over how their recorded interactions might be later used. For example, should practitioners be allowed to sample the recorded users' nonverbal behavior as references for transforming other unrelated recordings? Should practitioners be allowed to use the recorded avatar representations to generate new content? If so, how should the recorded users review the generated content and be properly credited and compensated? In addition, the model should also notify the new users of the possible nonverbal transformations and provide the original unchanged recording if requested by the new users. As further discussed in the next section, practitioners should develop additional tools, select real-world applications, and curate content that minimize unintended effects after nonverbal transformations.

6.3 Applications of Asynchronous VR Social Interactions

6.3.1 Content and Social Context of Recorded Social Interactions. When deploying these techniques into real-world scenarios, breaking the synchronization between an avatar's gaze and verbal behavior can distort the social context and original intentions of the recorded users. Because of this, practitioners should carefully deliberate on the choice of content and context of recorded social interactions such that the nonverbal transformations minimize the amount of out-of-context and inappropriate social behaviors (e.g., avatar looking at the new user when delivering an aggressive comment). For instance, rather than applying these transformations on personal or confrontational conversations, there may be greater benefits in transforming educational and academic discussions.

Another approach for mitigating these unintended effects is through developing tools and systems for practitioners and users to retrospectively label recording segments based on their suitability for nonverbal transformations. Given these labels, systems can then dynamically manipulate recordings at a finer granularity based on the content, social context, and user preferences. Relatedly, because virtual context and the time-dependence of human behavior such as synchrony and mimicry are closely related to social closeness, performance, and collaboration [57, 81], when transforming asynchronous social interactions, practitioners should also weigh the benefits of increased social presence and perceived attention against the potential downsides (e.g., decrease in task performance) associated with mismatched context and breaking natural synchrony. One possible remedy is to develop real-time post-processing techniques on transformed nonverbal behavior to preserve real-time synchrony and mimicry.

Additionally, as recorded users can take on individual roles (e.g., moderator, student) that are associated with different behavioral patterns, another approach for preserving the social context is to extend Wang et al.'s [84] approach by inserting gaze for recorded users through referencing existing gaze instances from similar contexts and users with matching roles. Finally, as with video sharing software such as YouTube and TikTok where content are curated for asynchronous viewing, we anticipate emerging content categories and curation guidelines that will help retain the relevance and significance of past VR social interactions when viewed out-of-context, at a later time, and after nonverbal transformations.

6.3.2 Real-world Scenarios of Asynchronous VR Social Interactions. As our transformations of nonverbal VR social interactions increased social presence and perceived attention, we envision several areas of real-world applications. One application area is in transforming immersive educational content. For example, we foresee educators leveraging nonverbal transformations to facilitate interactions between students and recorded discussion-based teaching content. As conducting lectures through teleconferencing software such as ZOOM can cause disengagement given the lack of nonverbal cues such as eye contact and upper-body movement [29], we see educators applying nonverbal transformations on recorded VR educational materials to increase students' sense of social presence and perceived attention. Additionally, these systems can enable students to participate in past lectures in more intuitive and naturalistic ways. For example, as the new student engages with the transformed asynchronous social interactions, future systems can allow them to record responses at different points of the session and adapt the shown interactions based on what the student had said (e.g., adding appropriate transitions, inserting gaze towards new users when they respond). Using these systems, instructors can later access and evaluate the students' responses.

More broadly speaking, transforming asynchronous social interactions allows users to be more integrated into past social interactions that may otherwise be inaccessible and feel remote to the new user. For instance, we imagine a future where aspiring writers can feel acclimated when observing creative discussions taking place within an experienced writer's room, up-and-coming artists can experience round table discussions with renowned artists, and social VR users in low-traffic time zones can be immersed in group interactions that are recorded during the peak hours of more populated time zones. Given these opportunities, we hope to foster broader discussions on how we can transform asynchronous social interactions in ethically and contextually appropriate ways across different applications.

6.4 Limitations

There are several limitations to our work. To start, as the Meta Quest 2 did not provide eye tracking, we used participants' head orientation to approximate gaze behavior. Using VR headsets with eye tracking capabilities would provide more accurate gaze information. In addition, although we compared different eye gaze threshold values when evaluating the robustness of our findings regarding gaze behavior, the lack of eye tracking data poses a limitation to our analysis. Furthermore, as our participants were recruited from a medium-sized private university and are therefore not representative of the general population, it is unclear how well our findings generalize towards the broader population. Similarly, we are also limited in our ability to draw practical implications across different user demographics such as age, gender, and ethnicity. Relatedly, because the VR recordings used for our study consisted of the same four avatars across stimuli, it is unclear whether our findings will generalize to transformed recordings of varying group sizes and demographic compositions such as age, race, and gender. As noted in Section 4.1, our sample size (N=128) fell short of our recruitment goal of 138. This limitation could have further hindered our ability to draw practical implications.

Furthermore, participants in our study completed questionnaires on a desktop computer after each VR session. While this procedure resembled real-world scenarios of users joining and leaving asynchronous VR social interactions, switching between VR and desktop PC when completing questionnaires can yield disorientation, disturbances, and undesirable effects on the variance of subjective responses [28, 64, 71]. It is also possible that these switches could negatively impact memory recall. Future studies can therefore consider distributing questionnaires in VR to reduce the undesirable effects changes in environments can have on questionnaire responses and further increase the reliability of our subjective measurements.

6.5 Future Directions

We applied transformation to recordings of VR group discussions with very little to no horizontal movement on the scale of meters, as would occur when walking around a room. While we found that certain manipulations increased metrics such as social presence, perceived attention and mutual gaze, it is unclear how these transformations affect recorded social interactions where avatars are actively moving around the virtual environment. Future work should therefore focus on developing and evaluating the implications of techniques that transform proxemics smoothly and consistently across time and are also reasonable within the spatial context. Relatedly, another direction for future research lies in contextualizing these techniques within the realm of virtual agents, where the recorded social interactions may be partially generated or accompanied by virtual agents. While we focused on nonverbal transformations, researchers can also design and evaluate transformations of verbal behaviors, for example through leveraging Large Language Models to alter recorded dialogues.

Further, although we manipulated head orientation as a proxy for gaze direction, future works can further decouple head and gaze direction to increase the perceived realism of recorded social interactions. For example, researchers can leverage VR headsets with eye tracking capabilities to reference both head and gaze motion in existing samples when inserting gaze. Similarly, rather than inserting gaze randomly and evenly across all recorded avatars, we imagine systems adding gaze more deliberately, for example inserting more gazes when avatars are speaking and sampling a large-scale dataset of natural social interactions to increase variability. While we set $r_{transformed}$ as 0.4 through early pilots and the amount of interpersonal distance manipulation based on Hall's proxemics Theory [31], additional evaluation should investigate how different levels of proxemics and gaze transformations affect the user's asynchronous VR experience. Relatedly, as avatar appearance is associated with self-presence [65], perception of others [25] as well as proximity and uncanniness [36], future studies should investigate how our findings such as the inverse relationship between interpersonal distance and perceived threat manifest during asynchronous interactions with different avatar representations and VR platforms.

Another avenue for future research is in applying these transformations on different types of social interactions. While we studied how individuals perceive transformed VR recordings of small group discussions with a guided prompt and measured content recall and persuasion, we envision systems employing our proposed transformations in conjunction with existing methods in asynchronous VR interaction [19, 23, 54] to video conferencing, learning and collaboration for both small and large groups, and consider metrics such as entitativity, learning transfer, and collaborative problem-solving.

7 CONCLUSION

VR is a powerful tool for facilitating asynchronous social interaction as it allows users to immersively experience past interactions in real time. Consequently, we envision future VR applications combining the convenience of asynchronous communication with the high social presence of VR. For example, these applications may allow students to join recorded discussion sections, remote employees to attend past group meetings, and social VR users in low-traffic time zones to experience crowds. While existing tools for recording and replaying social interactions in VR allow us to reconstruct past social interactions in high fidelity, we see opportunities for transforming recorded interactions to deliver an enhanced experience to the new user.

In this paper, we extended previous methods for transforming proxemics and gaze behavior of VR tracking data to assimilate new users into recorded group discussions. Through an evaluation of the proposed transformations, we found that the combination of spatial accommodation and inserting gaze increased social presence, perceived attention, and mutual gaze. Additionally, we found that closer interpersonal distance increased perceived threat. Based on our findings, we outlined actionable implications for practitioners such as educators and virtual meeting organizers to apply these techniques to real-world scenarios beyond prompt-guided group discussions.

ACKNOWLEDGMENTS

We would like to thank Cindy Chen, Cyan DeVeaux, Eugy Han, David Lindlbauer, and Yujie Tao for their valuable feedback. We thank Pedro Civita, Caroline Graham, and Nattakit Tankongchamruskul for their help on recording the group discussions and Jordan Egelman and Masaki Nakamura for coding recall question responses. We also thank Brian Beams for their help on user study logistics. Last, we would like to thank our reviewers and study participants. This work was supported by the Sozo Graduate Fellowship.

REFERENCES

[1] 2023. Global: VR headsets volume 2018-2028. https://www.statista.com/forecasts/1331896/vr-headset-sales-volume-worldwide. Accessed: 2023-8-16.

[2] Sean Andrist, Michael Gleicher, and Bilge Mutlu. 2017. Looking Coordinated: Bidirectional Gaze Mechanisms for Collaborative Interaction with Virtual Characters. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) *(CHI '17)*. Association for Computing Machinery, New York, NY, USA, 2571–2582. https://doi.org/10.1145/3025453.3026033

[3] Michael Argyle and Janet Dean. 1965. Eye-Contact, Distance and Affiliation. *Sociometry* 28, 3 (1965), 289–304. http://www.jstor.org/stable/2786027

[4] Michael Argyle, Roger Ingham, Florisse Alkema, and Margaret McCallin. 1973. The different functions of gaze. (1973).

[5] Sarune Baceviciute, Aske Mottelson, Thomas Terkildsen, and Guido Makransky. 2020. Investigating Representation of Text and Audio in Educational VR Using Learning Outcomes and EEG. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3313831.3376872

[6] Jeremy N. Bailenson, Andrew C. Beall, Jack Loomis, Jim Blascovich, and Matthew Turk. 2004. Transformed Social Interaction: Decoupling Representation from Behavior and Form in Collaborative Virtual Environments. *Presence: Teleoperators and Virtual Environments* 13, 4 (08 2004), 428–441. https://doi.org/10.1162/1054746041944803

[7] Jeremy N. Bailenson, Andrew C. Beall, Jack Loomis, Jim Blascovich, and Matthew Turk. 2005. Transformed Social Interaction, Augmented Gaze, and Social Influence in Immersive Virtual Environments. *Human Communication Research* 31, 4 (2005), 511–537. https://doi.org/10.1111/j.1468-2958.2005.tb00881.x

[8] Jeremy N. Bailenson, Jim Blascovich, Andrew C. Beall, and Jack M. Loomis. 2001. Equilibrium Theory Revisited: Mutual Gaze and Personal Space in Virtual Environments. *Presence: Teleoperators and Virtual Environments* 10, 6 (12 2001), 583–598. https://doi.org/10.1162/105474601753272844

[9] Jeremy N. Bailenson, Jim Blascovich, Andrew C. Beall, and Jack M. Loomis. 2003. Interpersonal Distance in Immersive Virtual Environments. *Personality and Social Psychology Bulletin* 29, 7 (2003), 819–833. https://doi.org/10.1177/0146167203029007002 PMID: 15018671.

[10] Jeremy N. Bailenson and Nick Yee. 2005. Digital Chameleons: Automatic Assimilation of Nonverbal Gestures in Immersive Virtual Environments. *Psychological Science* 16, 10 (2005), 814–819. https://doi.org/10.1111/j.1467-9280.2005.01619.x

[11] Tessa Beinema, Daniel Davison, Dennis Reidsma, Oresti Banos, Merijn Bruijnes, Brice Donval, Álvaro Fides Valero, Dirk Heylen, Dennis Hofs, Gerwin Huizing, Reshamashree B. Kantharaju, Randy Klaassen, Jan Kolkmeier, Kostas Konsolakis, Alison Pease, Catherine Pelachaud, Donatella Simonetti, Mark Snaith, Vicente Traver, Jorien van Loon, Jacky Visser, Marcel Weusthof, Fajrian Yunus, Hermie Hermens, and Harm op den Akker. 2021. Agents United: An Open Platform for Multi-Agent Conversational Systems. In *Proceedings of the 21st ACM International Conference on Intelligent Virtual Agents* (Virtual Event, Japan) *(IVA '21)*. Association for Computing Machinery, New York, NY, USA, 17–24. https://doi.org/10.1145/3472306.3478352

[12] Andrea Bönsch, Alexander R. Bluhm, Jonathan Ehret, and Torsten W. Kuhlen. 2020. Inferring a User's Intent on Joining or Passing by Social Groups. In *Proceedings of the 20th ACM International Conference on Intelligent Virtual Agents* (Virtual Event, Scotland, UK) *(IVA '20)*. Association for Computing Machinery, New York, NY, USA, Article 10, 8 pages. https://doi.org/10.1145/3383652.3423862

[13] Vojtěch Brůža, Jan Byška, Jan Mičan, and Barbora Kozlíková. 2021. VRdeo: Creating Engaging Educational Material for Asynchronous Student-Teacher Exchange Using Virtual Reality. *Comput. Graph.* 98, C (aug 2021), 280–292. https://doi.org/10.1016/j.cag.2021.06.009

[14] Gavin Buckingham. 2021. Hand Tracking for Immersive Virtual Reality: Opportunities and Challenges. *Frontiers in Virtual Reality* 2 (2021). https://doi.org/10.3389/frvir.2021.728461

[15] Alisa Burova, John Mäkelä, Hanna Heinonen, Paulina Becerril Palma, Jaakko Hakulinen, Viveka Opas, Sanni Siltanen, Roope Raisamo, and Markku Turunen. 2022. Asynchronous industrial collaboration: How virtual reality and virtual tools aid the process of maintenance method development and documentation creation. *Computers in Industry* 140 (2022), 103663. https://doi.org/10.1016/j.compind.2022.103663

[16] Fabio Buttussi and Luca Chittaro. 2018. Effects of Different Types of Virtual Reality Display on Presence and Learning in a Safety Training Scenario. *IEEE Transactions on Visualization and Computer Graphics* 24, 2 (2018), 1063–1076. https://doi.org/10.1109/TVCG.2017.2653117

[17] Scott Cardall, Edward Krupat, and Michael Ulrich. 2008. Live Lecture Versus Video-Recorded Lecture: Are Students Voting With Their Feet? *Academic medicine : journal of the Association of American Medical Colleges* 83 (01 2008), 1174–8. https://doi.org/10.1097/ACM.0b013e31818c6902

[18] Dorwin Cartwright. 1971. Risk taking by individuals and groups: An assessment of research employing choice dilemmas. *Journal of Personality and Social Psychology* 20, 3 (1971), 361.

[19] Kevin Chow, Caitlin Coyiuto, Cuong Nguyen, and Dongwook Yoon. 2019. Challenges and Design Considerations for Multimodal Asynchronous Collaboration in VR. *Proc. ACM Hum.-Comput. Interact.* 3, CSCW, Article 40 (nov 2019), 24 pages. https://doi.org/10.1145/3359142

[20] Jacob Cohen. 1992. A Power Primer. *Tutorials in Quantitative Methods for Psychology* 112 (07 1992). https://doi.org/10.1037/0033-2909.112.1.155

[21] James J. Cummings and Jeremy N. Bailenson. 2016. How Immersive Is Enough? A Meta-Analysis of the Effect of Immersive Technology on User Presence. *Media Psychology* 19, 2 (2016), 272–309. https://doi.org/10.1080/15213269.2015.1015740 arXiv:https://doi.org/10.1080/15213269.2015.1015740

[22] Géraldine Fauville, Anna C. Queiroz, Mufan Luo, Jeffrey Hancock, and Jeremy Bailenson. 2022. Impression Formation From Video Conference Screenshots: The Role of Gaze, Camera Distance, and Angle. *Technology, Mind, and Behavior* 3 (01 2022). https://doi.org/10.1037/tmb0000055

[23] Andreas Rene Fender and Christian Holz. 2022. Causality-Preserving Asynchronous Reality. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 634, 15 pages. https://doi.org/10.1145/3491102.3501836

[24] Tom Foulsham, Esther Walker, and Alan Kingstone. 2011. The where, what and when of gaze allocation in the lab and the natural environment. *Vision Research* 51, 17 (2011), 1920–1931. https://doi.org/10.1016/j.visres.2011.07.002

[25] Guo Freeman and Divine Maloney. 2021. Body, Avatar, and Me: The Presentation and Perception of Self in Social Virtual Reality. *Proc. ACM Hum.-Comput. Interact.* 4, CSCW3, Article 239 (jan 2021), 27 pages. https://doi.org/10.1145/3432938

[26] Doron Friedman, Anthony Steed, and Mel Slater. 2007. Spatial Social Behavior in Second Life. *Lecture Notes in Computer Science* 4722, 252–263. https://doi.org/10.1007/978-3-540-74997-4_23

[27] Andreas Girgensohn, Jennifer Marlow, Frank Shipman, and Lynn Wilcox. 2015. HyperMeeting: Supporting Asynchronous Meetings with Hypervideo. In *Proceedings of the 23rd ACM International Conference on Multimedia* (Brisbane, Australia) *(MM '15)*. Association for Computing Machinery, New York, NY, USA, 611–620. https://doi.org/10.1145/2733373.2806258

[28] Sarah Graf and Valentin Schwind. 2020. Inconsistencies of Presence Questionnaires in Virtual Reality. In *Proceedings of the 26th ACM Symposium on Virtual Reality Software and Technology* (Virtual Event, Canada) *(VRST '20)*. Association for Computing Machinery, New York, NY, USA, Article 60, 3 pages. https://doi.org/10.1145/3385956.3422105

[29] Chengyuan Jia Gulipari Maimaiti and Khe Foon Hew. 2021. Student disengagement in web-based videoconferencing supported online learning: an activity theory perspective. *Interactive Learning Environments* 0, 0 (2021), 1–20. https://doi.org/10.1080/10494820.2021.1984949

[30] Ihshan Gumilar, Amit Barde, Ashkan F. Hayati, Mark Billinghurst, Gun Lee, Abdul Momin, Charles Averill, and Arindam Dey. 2021. Connecting the Brains via Virtual Eyes: Eye-Gaze Directions and Inter-Brain Synchrony in VR. In *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI EA '21)*. Association for Computing Machinery, New York, NY, USA, Article 350, 7 pages. https://doi.org/10.1145/3411763.3451583

[31] E.T. Hall. 1966. *The Hidden Dimension*. Knopf Doubleday Publishing Group. https://books.google.ch/books?id=RdZOAAAAMAAJ

[32] Eugy Han, Mark R Miller, Cyan DeVeaux, Hanseul Jun, Kristine L Nowak, Jeffrey T Hancock, Nilam Ram, and Jeremy N Bailenson. 2023. People, places, and time: a large-scale, longitudinal study of transformed avatars and environmental context in group interaction in the metaverse. *Journal of Computer-Mediated Communication* 28, 2 (01 2023). https://doi.org/10.1093/jcmc/zmac031

[33] Chad Harms and Frank Biocca. 2004. Internal consistency and reliability of the networked minds measure of social presence. In *Seventh annual international workshop: Presence*, Vol. 2004. Universidad Politecnica de Valencia Valencia, Spain.

[34] Heiko Hecht, Robin Welsch, Jana Viehoff, and Matthew R. Longo. 2019. The shape of personal space. *Acta Psychologica* 193 (2019), 113–122. https://doi.org/10.1016/j.actpsy.2018.12.009

[35] Dirk Helbing and Péter Molnár. 1995. Social force model for pedestrian dynamics. *Phys. Rev. E* 51 (May 1995), 4282–4286. Issue 5. https://doi.org/10.1103/PhysRevE.51.4282

[36] Daniel Hepperle, Christian Felix Purps, Jonas Deuchler, and Matthias Wölfel. 2022. Aspects of Visual Avatar Appearance: Self-Representation, Display Type, and Uncanny Valley. *Vis. Comput.* 38, 4 (apr 2022), 1227–1244. https://doi.org/

10.1007/s00371-021-02151-0

[37] Valentin Holzwarth, Johannes Schneider, Joshua Handali, Joy Gisler, Christian Hirt, Andreas Kunz, and Jan vom Brocke. 2021. Towards estimating affective states in Virtual Reality based on behavioral data. *Virtual Reality* 25 (12 2021). https://doi.org/10.1007/s10055-021-00518-1

[38] Ann Huang, Pascal Knierim, Francesco Chiossi, Lewis L Chuang, and Robin Welsch. 2022. Proxemics for Human-Agent Interaction in Augmented Reality. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 421, 13 pages. https://doi.org/10.1145/3491102.3517593

[39] Wen Huang, Rod Roscoe, Mina Johnson-Glenberg, and Scotty Craig. 2020. Motivation, engagement, and performance across multiple virtual reality sessions and levels of immersion. *Journal of Computer Assisted Learning* 37 (12 2020). https://doi.org/10.1111/jcal.12520

[40] Alessandro Iop, Sahba Zojaji, and Christopher Peters. 2022. Don't Walk between Us: Adherence to Social Conventions When Joining a Small Conversational Group of Agents. In *Proceedings of the 22nd ACM International Conference on Intelligent Virtual Agents* (Faro, Portugal) *(IVA '22)*. Association for Computing Machinery, New York, NY, USA, Article 6, 8 pages. https://doi.org/10.1145/3514197.3549676

[41] Maidul Islam, Dan-A. Kim, and Minjoo Kwon. 2020. A Comparison of Two Forms of Instruction: Pre-Recorded Video Lectures vs. Live ZOOM Lectures for Education in the Business Management Field. *Sustainability* 12, 19 (2020). https://doi.org/10.3390/su12198149

[42] Dušan Jan and David R. Traum. 2005. Dialog Simulation for Background Characters. In *Intelligent Virtual Agents*, Themis Panayiotopoulos, Jonathan Gratch, Ruth Aylett, Daniel Ballin, Patricka Olivier, and Thomas Rist (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 65–74.

[43] Dusan Jan and David R. Traum. 2007. Dynamic Movement and Positioning of Embodied Agents in Multiparty Conversations. In *Proceedings of the 6th International Joint Conference on Autonomous Agents and Multiagent Systems* (Honolulu, Hawaii) *(AAMAS '07)*. Association for Computing Machinery, New York, NY, USA, Article 14, 3 pages. https://doi.org/10.1145/1329125.1329142

[44] Nathan Kogan and Michael A. Wallach. 1965. Risk Taking: A Study in Cognition and Personality.

[45] Nathan Kogan and Michael A Wallach. 1967. Risky-shift phenomenon in small decision-making groups: A test of the information-exchange hypothesis. *Journal of Experimental Social Psychology* 3, 1 (1967), 75–84.

[46] Jan Kolkmeier, Jered Vroon, and Dirk Heylen. 2016. Interacting with Virtual Agents in Shared Space: Single and Joint Effects of Gaze and Proxemics. In *Intelligent Virtual Agents*, David Traum, William Swartout, Peter Khooshabeh, Stefan Kopp, Stefan Scherer, and Anton Leuski (Eds.). Springer International Publishing, Cham, 1–14.

[47] Eric Krokos, Catherine Plaisant, and Amitabh Varshney. 2019. Virtual Memory Palaces: Immersion Aids Recall. *Virtual Real.* 23, 1 (mar 2019), 1–15. https://doi.org/10.1007/s10055-018-0346-3

[48] Claudia Kuster, Tiberiu Popa, Jean-Charles Bazin, Craig Gotsman, and Markus Gross. 2012. Gaze Correction for Home Video Conferencing. *ACM Trans. Graph.* 31, 6, Article 174 (nov 2012), 6 pages. https://doi.org/10.1145/2366145.2366193

[49] Helen E. Larkin. 2010. "But they won't come to lectures ..." The impact of audio recorded lectures on student experience and attendance. *Australasian Journal of Educational Technology* 26, 2 (2010). https://doi.org/10.14742/ajet.1093

[50] Kien Le. 2022. Pre-Recorded Lectures, Live Online Lectures, and Student Academic Achievement. *Sustainability* 14, 5 (2022). https://doi.org/10.3390/su14052910

[51] Jiwei Li, Will Monroe, Alan Ritter, Dan Jurafsky, Michel Galley, and Jianfeng Gao. 2016. Deep Reinforcement Learning for Dialogue Generation. In *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*, Jian Su, Kevin Duh, and Xavier Carreras (Eds.). Association for Computational Linguistics, Austin, Texas, 1192–1202. https://doi.org/10.18653/v1/D16-1127

[52] MultiMedia LLC. 2023. *Improve Human Connection in Video Conferences with NVIDIA Maxine Eye Contact.* https://developer.nvidia.com/blog/improve-human-connection-in-video-conferences-with-nvidia-maxine-eye-contact/

[53] Joan Llobera, Bernhard Spanlang, Giulio Ruffini, and Mel Slater. 2010. Proxemics with Multiple Dynamic Characters in an Immersive Virtual Environment. *ACM Trans. Appl. Percept.* 8, 1, Article 3 (nov 2010), 12 pages. https://doi.org/10.1145/1857893.1857896

[54] Thomas Lopez, Olivier Dumas, Fabien Danieau, Bertrand Leroy, Nicolas Mollet, and Jean-François Vial. 2017. A Playback Tool for Reviewing VR Experiences. In *Proceedings of the 23rd ACM Symposium on Virtual Reality Software and Technology* (Gothenburg, Sweden) *(VRST '17)*. Association for Computing Machinery, New York, NY, USA, Article 83, 2 pages. https://doi.org/10.1145/3139131.3141776

[55] Divine Maloney, Guo Freeman, and Donghee Yvette Wohn. 2020. "Talking without a Voice": Understanding Non-Verbal Communication in Social Virtual Reality. *Proc. ACM Hum.-Comput. Interact.* 4, CSCW2, Article 175 (oct 2020), 25 pages. https://doi.org/10.1145/3415246

[56] Mark Roman Miller, Cyan DeVeaux, Eugy Han, Nilam Ram, and Jeremy N. Bailenson. 2023. A Large-Scale Study of Proxemics and Gaze in Groups. In *2023 IEEE Conference Virtual Reality and 3D User Interfaces (VR)*. 409–417. https://doi.org/10.1109/VR55154.2023.00056

[57] Mark Roman Miller, Neeraj Sonalkar, Ade Mabogunje, Larry Leifer, and Jeremy Bailenson. 2021. Synchrony within Triads Using Virtual Reality. *Proc. ACM Hum.-Comput. Interact.* 5, CSCW2, Article 400 (oct 2021), 27 pages. https://doi.org/10.1145/3479544

[58] Mikhail Morozov, Alexey Gerasimov, and Mikhail Fominykh. 2012. vAcademia – Educational Virtual World with 3D Recording. In *2012 International Conference on Cyberworlds*. 199–206. https://doi.org/10.1109/CW.2012.35

[59] Mukesh Nathan, Mercan Topkara, Jennifer Lai, Shimei Pan, Steven Wood, Jeff Boston, and Loren Terveen. 2012. In Case You Missed It: Benefits of Attendee-Shared Annotations for Non-Attendees of Remote Meetings. In *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work* (Seattle, Washington, USA) *(CSCW '12)*. Association for Computing Machinery, New York, NY, USA, 339–348. https://doi.org/10.1145/2145204.2145259

[60] Lasse T. Nielsen, Matias B. Møller, Sune D. Hartmeyer, Troels C. M. Ljung, Niels C. Nilsson, Rolf Nordahl, and Stefania Serafin. 2016. Missing the Point: An Exploration of How to Guide Users' Attention during Cinematic Virtual Reality. In *Proceedings of the 22nd ACM Conference on Virtual Reality Software and Technology* (Munich, Germany) *(VRST '16)*. Association for Computing Machinery, New York, NY, USA, 229–232. https://doi.org/10.1145/2993369.2993405

[61] David Novick, Laura J. Hinojos, Aaron E. Rodriguez, Adriana Camacho, and Mahdokht Afravi. 2018. Conversational Interaction with Multiple Agents Initiated via Proxemics and Gaze. In *Proceedings of the 6th International Conference on Human-Agent Interaction* (Southampton, United Kingdom) *(HAI '18)*. Association for Computing Machinery, New York, NY, USA, 356–358. https://doi.org/10.1145/3284432.3287185

[62] Soo Youn Oh, Jeremy Bailenson, Erika Weisz, and Jamil Zaki. 2016. Virtually old: Embodied perspective taking and the reduction of ageism under threat. *Computers in Human Behavior* 60 (2016), 398–410. https://doi.org/10.1016/j.chb.2016.02.007

[63] Barbara C Panther, Jennifer A Mosse, and Wendy Wright. 2011. Recorded lectures don't replace the 'real thing': What the students say. In *Proceedings of The Australian Conference on Science and Mathematics Education*. 127–132.

[64] Susanne Putze, Dmitry Alexandrovsky, Felix Putze, Sebastian Höffner, Jan David Smeddinck, and Rainer Malaka. 2020. Breaking The Experience: Effects of Questionnaires in VR User Studies. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–15. https://doi.org/10.1145/3313831.3376144

[65] Rabindra Ratan. 2013. Self-presence, explicated: Body, emotion, and identity extension into the virtual self. In *Handbook of research on technoself: Identity in a technological society*. IGI Global, 322–336.

[66] Frank Rodriguez. 2017. Hot and Cold: Heatmaps in VR. *The Guardian* (2017). https://blog.youtube/news-and-events/hot-and-cold-heatmaps-in-vr/

[67] Daniel Roth, Gary Bente, Peter Kullmann, David Mal, Chris Felix Purps, Kai Vogeley, and Marc Erich Latoschik. 2019. Technologies for Social Augmentations in User-Embodied Virtual Reality. In *Proceedings of the 25th ACM Symposium on Virtual Reality Software and Technology* (Parramatta, NSW, Australia) *(VRST '19)*. Association for Computing Machinery, New York, NY, USA, Article 5, 12 pages. https://doi.org/10.1145/3359996.3364269

[68] Daniel Roth, Constantin Klelnbeck, Tobias Feigl, Christopher Mutschler, and Marc Erich Latoschik. 2018. Beyond Replication: Augmenting Social Behaviors in Multi-User Virtual Realities. In *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. 215–222. https://doi.org/10.1109/VR.2018.8447550

[69] Daniel Roth, Peter Kullmann, Gary Bente, Dominik Gall, and Marc Erich Latoschik. 2018. Effects of Hybrid and Synthetic Social Gaze in Avatar-Mediated Interactions. In *2018 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*. 103–108. https://doi.org/10.1109/ISMAR-Adjunct.2018.00044

[70] Florian Schier, Krishnan Chandran, and Matthew McGinity. 2022. TeachInVR: A virtual reality classroom for remote education. In *2022 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*. 283–286. https://doi.org/10.1109/VRW55335.2022.00064

[71] Valentin Schwind, Pascal Knierim, Nico Haas, and Niels Henze. 2019. Using Presence Questionnaires in Virtual Reality. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3290605.3300590

[72] Yangming Shi, Jing Du, Changbum R. Ahn, and Eric Ragan. 2019. Impact assessment of reinforced learning methods on construction workers' fall risk behavior using virtual reality. *Automation in Construction* 104 (2019), 197–214. https://doi.org/10.1016/j.autcon.2019.04.015

[73] Mel Slater, Amela Sadagic, Martin Usoh, and Ralph Schroeder. 2000. Small-Group Behavior in a Virtual and Real Environment: A Comparative Study. *Presence: Teleoper. Virtual Environ.* 9, 1 (feb 2000), 37–51. https://doi.org/10.1162/105474600566600

[74] Mel Slater, Anthony Steed, John McCarthy, and Francesco Maringelli. 1998. The Influence of Body Movement on Subjective Presence in Virtual Environments. *Human Factors* 40, 3 (1998), 469–477. https://doi.org/10.1518/001872098779591368

[75] Mel Slater and Sylvia Wilbur. 1997. A Framework for Immersive Virtual Environments Five: Speculations on the Role of Presence in Virtual Environments. *Presence: Teleoper. Virtual Environ.* 6, 6 (dec 1997), 603–616. https://doi.org/10.1162/pres.1997.6.6.603

[76] Yoana Slavova and Mu Mu. 2018. A Comparative Study of the Learning Outcomes and Experience of VR in Education. In *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. 685–686. https://doi.org/10.1109/VR.2018.8446486

[77] Harrison Jesse Smith and Michael Neff. 2018. Communication Behavior in Embodied Virtual Reality. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI ’18)*. Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3173574.3173863

[78] Agnieszka Sorokowska, Piotr Sorokowski, Peter Hilpert, Katarzyna Cantarero, Tomasz Frackowiak, Khodabakhsh Ahmadi, Ahmad M. Alghraibeh, Richmond Aryeetey, Anna Bertoni, Karim Bettache, Sheyla Blumen, Marta Błażejewska, Tiago Bortolini, Marina Butovskaya, Felipe Nalon Castro, Hakan Cetinkaya, Diana Cunha, Daniel David, Oana A. David, Fahd A. Dileym, Alejandra del Carmen Domínguez Espinosa, Silvia Donato, Daria Dronova, Seda Dural, Jitka Fialová, Maryanne Fisher, Evrim Gulbetekin, Aslıhan Hamamcıoğlu Akkaya, Ivana Hromatko, Raffaella Iafrate, Mariana Iesyp, Bawo James, Jelena Jaranovic, Feng Jiang, Charles Obadiah Kimamo, Grete Kjelvik, Fırat Koç, Amos Laar, Fívia de Araújo Lopes, Guillermo Macbeth, Nicole M. Marcano, Rocio Martinez, Norbert Mesko, Natalya Molodovskaya, Khadijeh Moradi, Zahrasadat Motahari, Alexandra Mühlhauser, Jean Carlos Natividade, Joseph Ntayi, Elisabeth Oberzaucher, Oluyinka Ojedokun, Mohd Sofian Bin Omar-Fauzee, Ike E. Onyishi, Anna Paluszak, Alda Portugal, Eugenia Razumiejczyk, Anu Realo, Ana Paula Relvas, Maria Rivas, Muhammad Rizwan, Svjetlana Salkičević, Ivan Sarmány-Schuller, Susanne Schmehl, Oksana Senyk, Charlotte Sinding, Eftychia Stamkou, Stanislava Stoyanova, Denisa Šukolová, Nina Sutresna, Meri Tadinac, Andero Teras, Edna Lúcia Tinoco Ponciano, Ritu Tripathi, Nachiketa Tripathi, Mamta Tripathi, Olja Uhryn, Maria Emília Yamamoto, Gyesook Yoo, and Jr. John D. Pierce. 2017. Preferred Interpersonal Distances: A Global Comparison. *Journal of Cross-Cultural Psychology* 48, 4 (2017), 577–592. https://doi.org/10.1177/0022022117698039 arXiv:https://doi.org/10.1177/0022022117698039

[79] Kota Takahashi, Yasuyuki Inoue, and Michiteru Kitazaki. 2022. Interpersonal Distance to a Speaking Avatar: Loudness Matters Irrespective of Contents. In *2022 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*. 774–775. https://doi.org/10.1109/VRW55335.2022.00238

[80] Theresa Jean Tanenbaum, Nazely Hartoonian, and Jeffrey Bryan. 2020. "How Do I Make This Thing Smile?": An Inventory of Expressive Nonverbal Communication in Commercial Social Virtual Reality Platforms. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI ’20)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3313831.3376606

[81] B. Tarr, M. Slater, and E. Cohen. 2018. Synchrony and social connection in immersive Virtual Reality. *Scientific Reports* 8, 1 (27 Feb 2018), 3693. https://doi.org/10.1038/s41598-018-21765-4

[82] Roel Vertegaal, Robert Slagter, Gerrit van der Veer, and Anton Nijholt. 2001. Eye Gaze Patterns in Conversations: There is More to Conversational Agents than Meets the Eyes. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Seattle, Washington, USA) *(CHI ’01)*. Association for Computing Machinery, New York, NY, USA, 301–308. https://doi.org/10.1145/365024.365119

[83] Cheng Yao Wang, Mose Sakashita, Upol Ehsan, Jingjin Li, and Andrea Stevenson Won. 2020. Again, Together: Socially Reliving Virtual Reality Experiences When Separated. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI ’20)*. Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3313831.3376642

[84] Portia Wang, Mark R. Miller, and Jeremy N. Bailenson. 2023. The Belated Guest: Exploring the Design Space for Transforming Asynchronous Social Interactions in Virtual Reality. In *2023 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*. 617–618. https://doi.org/10.1109/VRW58643.2023.00151

[85] Robin Welsch, Christoph von Castell, Martin Rettenberger, Daniel Turner, Heiko Hecht, and Peter Fromberger. 2020. Sexual attraction modulates interpersonal distance and approach-avoidance movements towards virtual agents in males. *PLOS ONE* 15, 4 (04 2020), 1–18. https://doi.org/10.1371/journal.pone.0231539

[86] Julie Williamson, Jie Li, Vinoba Vinayagamoorthy, David A. Shamma, and Pablo Cesar. 2021. Proxemics and Social Interactions in an Instrumented Virtual Reality Workshop. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI ’21)*. Association for Computing Machinery, New York, NY, USA, Article 253, 13 pages. https://doi.org/10.1145/3411764.3445729

[87] Andrea Won, Brian Perone, Michelle Friend, and Jeremy Bailenson. 2016. Identifying Anxiety Through Tracked Head Movements in a Virtual Classroom. *Cyberpsychology, behavior and social networking* 19 (06 2016), 380–387. https://doi.org/10.1089/cyber.2015.0326

[88] Jason Wu, Sayan Ghosh, Mathieu Chollet, Steven Ly, Sharon Mozgai, and Stefan Scherer. 2018. NADiA: Neural Network Driven Virtual Human Conversation Agents. In *Proceedings of the 18th International Conference on Intelligent Virtual Agents* (Sydney, NSW, Australia) *(IVA ’18)*. Association for Computing Machinery, New York, NY, USA, 173–178. https://doi.org/10.1145/3267851.3267860

[89] Weilai Xu, Fred Charles, and Charlie Hargood. 2023. Generating Stylistic and Personalized Dialogues for Virtual Agents in Narratives. In *Proceedings of the 2023 International Conference on Autonomous Agents and Multiagent Systems* (London, United Kingdom) *(AAMAS ’23)*. International Foundation for Autonomous Agents and Multiagent Systems, Richland, SC, 737–746.

[90] Fangkai Yang and Christopher Peters. 2019. App-LSTM: Data-Driven Generation of Socially Acceptable Trajectories for Approaching Small Groups of Agents. In *Proceedings of the 7th International Conference on Human-Agent Interaction* (Kyoto, Japan) *(HAI ’19)*. Association for Computing Machinery, New York, NY, USA, 144–152. https://doi.org/10.1145/3349537.3351885

[91] Nick Yee, Jeremy Bailenson, Mark Urbanek, Francis Chang, and Dan Merget. 2007. The Unbearable Likeness of Being Digital: The Persistence of Nonverbal Social Norms in Online Virtual Environments. *Cyberpsychology & behavior: the impact of the Internet, multimedia and virtual reality on behavior and society* 10 (03 2007), 115–21. https://doi.org/10.1089/cpb.2006.9984

[92] Pavel Zahorik and Rick L. Jenison. 1998. Presence as Being-in-the-World. *Presence: Teleoperators and Virtual Environments* 7, 1 (02 1998), 78–89. https://doi.org/10.1162/105474698565541

[93] C.A. Zanbaka, B.C. Lok, S.V. Babu, A.C. Ulinski, and L.F. Hodges. 2005. Comparison of path visualizations and cognitive measures relative to travel technique in a virtual environment. *IEEE Transactions on Visualization and Computer Graphics* 11, 6 (2005), 694–705. https://doi.org/10.1109/TVCG.2005.92

[94] Johannes A. Zuber, Helmut W. Crott, and Joachim Werner. 1992. Choice shift and group polarization : an analysis of the status of arguments and social decision schemes. *Journal of Personality and Social Psychology* 62 (1992), 50–61. https://doi.org/10.1037/0022-3514.62.1.50