Body Crumple, Sound Intrusion, and Embodiment Violation: Toward a Framework for Miscommunication in VR

Daniel Akselrad
Stanford University
Stanford, CA, USA
daniel.akselrad@stanford.edu

Cyan DeVeaux
Stanford University
Stanford, CA, USA
cyanjd@stanford.edu

Eugy Han
Stanford University
Stanford, CA, USA
eugyoung@stanford.edu

Mark Roman Miller
Stanford University
Stanford, CA, USA
mrmillr@stanford.edu

Jeremy N. Bailenson
Stanford University
Stanford, CA, USA
bailenso@stanford.edu

ABSTRACT

The advent of widely-accessible VR has enabled individuals to communicate—and miscommunicate—in new ways. To explore these miscommunications, we introduce a preliminary framework based on events that occurred during 3600 minutes of observation inside a university course taught in VR. During the course, 250 people met in groups for about 20 minutes per session. We identify three types of miscommunication that routinely occurred: body crumple, sound intrusion, and embodiment violation. By mapping the affordances by which social VR fails to facilitate effective communication, we hope to provide educators, developers, and virtual ethnographers with a means for understanding and navigating the challenges of VR-based collaboration.

CCS CONCEPTS

• **Human-Centered Computing**; • **Collaborative and Social Computing**; • **Web-based interaction**;

KEYWORDS

Social VR; Interactive Learning Environments; Empirical Studies in Collaborative and Social Computing; Education

ACM Reference Format:
Daniel Akselrad, Cyan DeVeaux, Eugy Han, Mark Roman Miller, and Jeremy N. Bailenson. 2023. Body Crumple, Sound Intrusion, and Embodiment Violation: Toward a Framework for Miscommunication in VR. In *Computer Supported Cooperative Work and Social Computing (CSCW '23 Companion), October 14–18, 2023, Minneapolis, MN, USA.* ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3584931.3606968

1 INTRODUCTION

New ways of seeing, communicating, and embodying social space in virtual reality (VR) have brought forth novel possibilities for interpersonal miscommunication. The stakes of VR miscommunication are high. Now that universities, corporations, and governments rely on social VR platforms to facilitate remote work, it has become necessary to understand where and how communication breaks down. This means addressing the types of messages and behaviors that emerging immersive technologies cannot reliably accommodate. To better understand where VR fails to facilitate effective communication between users, we observed 250 students interacting across 180 20-minute formal discussion sessions convened as part of a Stanford University course taught in VR. Grounded on our prior work on group interaction in the metaverse [6], this paper introduces a preliminary framework identifying three types of failures that occur during interactions in social VR.

Previous scholars in CSCW and HCI have outlined theoretical frameworks for understanding miscommunications that occur in VR. Loomis' concept of "mixed attribution," for example, is instructive here [13], as is Slater's prescient warning concerning the dangers of "superrealism" [16]. Others have correlated perceived violations of personal space, physical-virtual collisions, and VR safety risks with a drop in communication quality [10, 11, 18, 20]. However, most examinations of miscommunication due to failures in social VR have stemmed from small laboratory studies, anecdotal accounts, or analysis of existing, curated content on platforms such as YouTube or Reddit [4, 7]. This paper is the first to arise from an examination of specific VR miscommunication events, and to address them in the context of months-long periods of repeated interaction inside a single immersive VR platform.

2 METHODS

Our data is a subset of the Stanford Longitudinal VR Classroom Dataset (SLVRClaD), which includes data from approximately 500 students across two years, each of whom took one of four separate courses taught using the ENGAGE social VR platform accessed via Meta Quest 2 VR headsets. Students met in groups ranging from two to 40, and consented to have their verbal, nonverbal, and performance data continually tracked during each course, typically for about eight weekly sessions which lasted about 20 minutes per session. In addition, each student self-reported about their experience after each session. This paper utilizes previously unreported data from the dataset, and focuses on two separate courses, including 250 participants over 3600 minutes across 180 sessions.

CSCW '23 Companion, October 14–18, 2023, Minneapolis, MN, USA

ACM ISBN 979-8-4007-0129-0/23/10…$15.00
https://doi.org/10.1145/3584931.3606968

There were two stages of analyses: first, to conduct observations, the first author reviewed 180 discussion sessions after they had taken place, using ENGAGE MyRecording files on a desktop computer (.myrec). Because these files capture all the tracking and event data from VR sessions including avatars, movements, audio, and objects in 3D space, the coders were able to examine and record interactions from any angle or distance. It was even possible while embodying an avatar to experience recordings of these discussion sessions from any location and angle within the scene, including from within participants' avatars' heads, behind their teeth and eyes (See Figure 2). The first author captured video recordings, screenshots, and field notes which were then used to track miscommunication events. Together, the authors of the study generated this taxonomy after observing which types of events occurred most frequently (more than 20 times) throughout the term.

Following the first stage of analysis, a secondary coder investigated a subset of the data (n=48 sessions) to confirm the prevalence of the miscommunication events. For each discussion session, the coder identified events in terms of our preliminary taxonomy of the three event types, examining whether these events did or did not occur during a given session on a binary scale.

3 A FRAMEWORK FOR SOCIAL VR FAILURES

We proceed from the standpoint that to understand VR, one needs to examine the coordination of tracking, rendering, and display. Together, each of these technical parameters is required to produce a persistent perceptual environment for the user. The failure of any one of these factors to accommodate the affordances of face-to-face (FtF) communication affects users' abilities to carry out goal-oriented interactions in the Immersive Virtual Environment (IVE). Insofar as accidents reveal something of the substance of new media technologies, it is by studying the mishaps, failures, and the uncanny incongruities of VR that we can better understand the medium as it exists in its current state. For purposes of this study, we operationally define miscommunication as the inability for senders and receivers to faithfully represent themselves, their verbal messages, or their paralinguistic communications as a result of the medium's affordances. When an IVE fails to translate the very voices and behaviors it is designed to support (or fails to preclude the very mishaps it is designed to prevent), this can lead to a breakdown in effective communication between participants.

3.1 Body Crumple

Participants encountered difficulties controlling their avatars' bodies in virtual space, often resulting in a new type of nonverbal miscommunication: body crumple. These were instances where the inside-out Simultaneous Localization and Mapping in their Quest 2 headsets failed to translate a user's body movements onto an avatar's position as a possible configuration within the social VR platform [2]. Flawed inverse kinematics often resulted in bodies appearing to crumple to the floor and contort themselves into anatomically impossible positions (See Figure 1). While various platforms handle behaviorally unrealistic configurations of head and hand controllers differently (whether due to tracking errors or users placing the hand controllers far away from the headsets), there is no perfect solution to this in terms of avatar design. Amidst a paucity of nonverbal cues such as facial expression and social context cues, these crumpling events complicated participants' abilities to infer whether their communication partners were present, aware, and ready to communicate. When caught off guard by their partners' sudden paralinguistic contortions, collaborators became unsure of how to respond, often waiting seconds, and sometimes even minutes, for an indication that their partner was still present in the IVE. Across 48 sessions that were evaluated by the second coder, there were a total of 75 instances of body crumple (on an average session, M=1.56, SD=1.86, min=0, max=7).

Figure 1: Two instances of body crumple. In each panel, a student has unknowingly and unintentionally caused their avatar to crumple and contort into an anatomically impossible position during a discussion session.

3.2 Sound Intrusion

Miscommunication occurred when audible sounds crafted to convey meaning in another medium, or ambient noise from another source, seeped into the IVE through a participant's headset microphone. These sound intrusions occurred, for example, when a participant's alarm clock rang out, when a fire engine siren passed a participant's bedroom, or when other nearby "real-world distractions" were taking place [14]. This audible interference that the IVE was not equipped to filter out introduced irrelevant, confounding stimuli to an already complex perceptual illusion. As one participant put it, "There were quite some distractions today, with strange sounds coming from unidentifiable sources." Often, these events induced what Steinecke and Bruder call a profoundly confusing "place and plausibility illusion," whereby users became unsure of their surroundings, due to the mixing of artifacts between virtual and real-world environments [17]. At times, these sounds seemed to be a relevant part of the IVE, prompting participants to switch tasks to search for the sound's origin. In several extreme cases, voices from a nearby technical support Zoom call were loud enough that a participant's microphone registered them as voice, causing their avatar's mouth to move along with the instructor's words (in an unwitting act of ventriloquism). In these extreme cases, these intrusions of the physical world into the IVE were more than mere background

noise over a telephone wire—they violated communication theorist H. P. Grice's maxim of relation ('Be relevant') [5] because they arrived through the animated mouths of avatars, taking the form of voice. In cases where these sounds distracted participants, these intrusions reduced the extent of users' presence in the IVE—"the perceptual illusion of nonmediation" that makes VR maximally immersive [12]. Participants inside the IVE who were within earshot of these intrusions amplified by the microphone were left to infer the origin, the purpose, and the intentionality of this extraneous information seeping in. Across 48 sessions that were evaluated by the second coder, there were a total of 32 instances of sound intrusion (on an average session, M=0.67, SD=1.26, min=0, max=6).

3.3 Embodiment Violation

The physical world has not prepared people to inhabit the body space of another individual—especially that of a classmate, a colleague, or an educator. Yet some social VR platforms make this eerily possible. Whereas Blackwell et al. have addressed the role of intentional personal space violations in virtual harassment [1], in our case students spawning into an IVE discussion session unintentionally entered other students' body space without even knowing it. Because ENGAGE enables multiple users to spawn into the same location by default, these participants traveled through other students' virtual bodies by mistake, allowing them to see "inside" one another's body (e.g., teeth, eyes, neck). Most often, participants responded to these violations on a significant delay, moving out of position only after realizing their avatars' limbs and torso were entangled in those of others. In extreme cases, participants carried on with speaking to other group members, not realizing that their avatar was tangled up in the body of another. Such scenes, had they been physical, could understandably be recognized as inappropriate touching in the work environment. Embodiment violations of what is socially understood to be another person's proxemic boundaries will require new considerations for norms of bodily privacy [11] (See Figure 2). Across 48 sessions that were evaluated by the second coder, there were a total of 52 instances of embodiment violation (on an average session, M=1.08, SD=1.61, min=0, max=6).

Figure 2: Two instances of embodiment violation. In the left panel, three avatars have collided and are sharing body space, unbeknownst to the users controlling them. The right panel depicts the view from inside a participant's avatar's head during a discussion session.

4 DISCUSSION

Where body crumple, sound intrusion, and embodiment violation occurred, each event presented challenges to social norms of bodily comportment established in a largely FtF environment. Prominent marketing of VR technology champions the medium's ability to facilitate unprecedented degrees of selective self-presentation: to control one's bodily movements and appearance, or to walk in the shoes of others. It is worth noting, then, that the participants we observed experiencing these three types of miscommunications in no way intended to experience them, and in many cases, they did not realize they had taken part in causing them. On the bright side, because communication over time is iterative and tends toward increasing complexity [3], social VR collaborators have opportunities to work together to get past these mishaps that otherwise hinder the quality of their interactions. Often where these three types of miscommunication occurred, participants evidenced spontaneous ingenuity to see their way to solving the problem. Still, as universities and corporations turn to social VR, and as governments move toward embracing the metaverse as the site of virtual embassies or "metaverse nations" [9, 19], the inability to control one's selective self-presentation will factor in determining whether individuals will be able to learn, to negotiate, or to access government services. Imagine, for example, being unable to plead your case for asylum in a virtual embassy, unintentionally violating cultural norms of behavior, or having your virtual body space violated during a meeting: wider adoption of social VR will raise the stakes of VR miscommunication.

This study focuses on these communications as they have been recorded in VR, highlighting the virtual side of these interactions. Therefore, a limitation of this work is its exclusion of VR-bystander dynamics and miscommunications that have begun to be explored in other work [15]. Another limitation of this study concerns the qualitative criteria for what constitutes a single occurrence of each of these events. Body crumples can be observed to have occurred more frequently if one considers partial crumpling—events in which an avatar's arms crumple but not its legs, for example. Similarly, embodiment violations (pictured above) may also include the many instances in which a person violated their own body space (e.g., inserting an arm directly through one's own avatar's torso), as well as events marked by repeated collisions of several avatars. Time, too, is a variable: it is possible that over time, users became more experienced and caused fewer of these events. To determine this would entail further investigation of discussion sessions held later in the term. Logging and coding these events in real time would enable the implementation of an additional survey that could ascertain each user's reactions to these events through self reporting. To that end, now that we know these types of miscommunication exist, future work can include a qualitative self-report survey by which to ascertain how these failures of social VR affected trust,

awareness, and copresence between collaborating participants. Importantly, to observe these events in real time would risk incurring observation effects that might impact the way these events play out between collaborators. Another constraint is that this investigation was limited to interactions in ENGAGE. However, each of the miscommunication events named here has the potential to occur on different social VR platforms: users of the platform VRChat, for example, have reported instances of "full body tracking glitches" on online forums [8].

In future work, we plan to complete evaluation of the remaining subset of sessions to better understand the prevalence of each miscommunication event across time and groups. Moreover, we plan to have additional coders evaluate this data to ensure reliability of the ratings. It is possible that there are other types of VR miscommunication events taking place on different timescales inside IVEs. More work is needed to discover these instances where social VR fails to facilitate interpersonal communication and selective self-presentation, to document them, and to continue building a framework by which to understand how and why these events arise.

5 CONCLUSION

The success of VR technologies to adapt established social norms of in-person collaboration will depend on the medium's ability to persistently support verbal and nonverbal cues. To the extent that 'code is law' in virtual space, developers of social VR should take note and take action to protect users by programming their IVEs to make it impossible to encounter the types of miscommunication we identify here [1]. For educators and virtual ethnographers, these findings suggest that while social VR presents a compelling model for teaching, learning, and participatory observation, it will become crucial to understand the pitfalls of VR miscommunication in order to avoid them going forward. Participants were eager to tell us after their first week of discussion sessions that they felt "the possibilities are limitless" for VR: it is up to the field to envision ways of protecting collaborators' implicit expectations of social space, privacy, and bodily autonomy to safeguard both the communication quality and ethical integrity of virtual environments.

ACKNOWLEDGMENTS

The authors wish to thank Brian Beams, Tobin Asher, and Mark York for their support with the course, and Brian Park for his work coding .myrec files.

REFERENCES

[1] Lindsay Blackwell, Nicole Ellison, Natasha Elliott-Deflo, and Raz Schwartz. 2019. Harassment in social VR: Implications for design. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE, 854–855.

[2] Polona Caserman, Philipp Achenbach, and Stefan Göbel. 2019. Analysis of inverse kinematics solutions for full-body reconstruction in virtual reality. In *2019 IEEE 7th International Conference on Serious Games and Applications for Health (SeGAH)*. IEEE, 1–8.

[3] Frank EX Dance. 1967. A helical model of communication. *Human communication theory* (1967), 294–98.

[4] Emily Dao, Andreea Muresan, Kasper Hornbæk, and Jarrod Knibbe. 2021. Bad breakdowns, useful seams, and face slapping: Analysis of vr fails on youtube. In *Proceedings of the 2021 chi conference on human factors in computing systems*. 1–14.

[5] Herbert P Grice. 1975. Logic and conversation. In *Speech acts*. Brill, 41–58.

[6] 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 (2023), zmac031.

[7] Markus Jelonek. 2023. VRtoER: When Virtual Reality leads to Accidents: A Community on Reddit as Lens to Insights about VR Safety. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–6.

[8] Kagumura. 2021. Full body glitching. https://www.reddit.com/r/VRchat/comments/qlw3uj/full_body_glitching/

[9] Nick Kelly and Marcus Foth. 2022. An entire Pacific country will upload itself to the metaverse. it's a desperate plan – with a hidden message. https://theconversation.com/an-entire-pacific-country-will-upload-itself-to-the-metaverse-its-a-desperate-plan-with-a-hidden-message-194728

[10] Kangsoo Kim, Divine Maloney, Gerd Bruder, Jeremy N Bailenson, and Gregory F Welch. 2017. The effects of virtual human's spatial and behavioral coherence with physical objects on social presence in AR. *Computer Animation and Virtual Worlds* 28, 3-4 (2017), e1771.

[11] Anya Kolesnichenko, Joshua McVeigh-Schultz, and Katherine Isbister. 2019. Understanding emerging design practices for avatar systems in the commercial social vr ecology. In *Proceedings of the 2019 on Designing Interactive Systems Conference*. 241–252.

[12] Matthew Lombard and Theresa Ditton. 1997. At the heart of it all: The concept of presence. *Journal of computer-mediated communication* 3, 2 (1997), JCMC321.

[13] Jack M Loomis. 1992. Presence and distal attribution: Phenomenology, determinants, and assessment. In *Human Vision, Visual Processing, and Digital Display III*, Vol. 1666. SPIE, 590–595.

[14] Catherine Oh, Fernanda Herrera, and Jeremy Bailenson. 2019. The effects of immersion and real-world distractions on virtual social interactions. *Cyberpsychology, Behavior, and Social Networking* 22, 6 (2019), 365–372.

[15] Joseph O'Hagan, Julie R Williamson, Florian Mathis, Mohamed Khamis, and Mark McGill. 2023. Re-evaluating vr user awareness needs during bystander interactions. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–17.

[16] Mel Slater, Cristina Gonzalez-Liencres, Patrick Haggard, Charlotte Vinkers, Rebecca Gregory-Clarke, Steve Jelley, Zillah Watson, Graham Breen, Raz Schwarz, William Steptoe, et al. 2020. The ethics of realism in virtual and augmented reality. *Frontiers in Virtual Reality* 1 (2020), 1.

[17] Frank Steinicke and Gerd Bruder. 2014. A self-experimentation report about long-term use of fully-immersive technology. In *Proceedings of the 2nd ACM symposium on Spatial user interaction*. 66–69.

[18] Kurt Thomas, Devdatta Akhawe, Michael Bailey, Dan Boneh, Elie Bursztein, Sunny Consolvo, Nicola Dell, Zakir Durumeric, Patrick Gage Kelley, Deepak Kumar, et al. 2021. Sok: Hate, harassment, and the changing landscape of online abuse. In *2021 IEEE Symposium on Security and Privacy (SP)*. IEEE, 247–267.

[19] Jim Wyss. 2021. Barbados Embassy near you could soon be the one on Metaverse. https://www.bloomberg.com/news/articles/2021-12-14/barbados-tries-digital-diplomacy-with-planned-metaverse-embassy#xj4y7vzkg

[20] Qingxiao Zheng, Shengyang Xu, Lingqing Wang, Yiliu Tang, Rohan C Salvi, Guo Freeman, and Yun Huang. 2023. Understanding Safety Risks and Safety Design in Social VR Environments. *Proceedings of the ACM on Human-Computer Interaction* 7, CSCW1 (2023), 1–37.