Social Interactions in Virtual Reality: What Cues Do People Use Most and How

Jingchao Fang
jcfang@ucdavis.edu
University of California, Davis
Davis, California, USA

Victoria Chang
vchang90@umd.edu
University of Maryland, College Park
College Park, Maryland, USA

Ge Gao
gegao@umd.edu
University of Maryland, College Park
College Park, Maryland, USA

Hao-Chuan Wang
hciwang@ucdavis.edu
University of California, Davis
Davis, California, USA

ABSTRACT

Social virtual reality (VR) has become increasingly popular. People today can leverage VR to meet and interact with others, especially when face-to-face (F2F) social events are not as practical. While strategies and cues being used in F2F interactions have been previously studied in CSCW, it remains unclear how people use various cues to interact with others in VR-based events, and whether the usage of different cues in such VR events differs from F2F interactions. To start addressing these questions, we conducted an online survey study, inviting VR event attendees to recall their recent social interaction experiences and cues usage in initiating, maintaining, and finishing interactions with event co-attendees in VR and F2F situations respectively (N=51). Results from the survey revealed differential cues usage in VR-based and in F2F interactions, which could shed light on the design of VR tools toward supporting online social events.

1 INTRODUCTION

The human condition of being naturally social [7] has led to the needs of communication tools that assist our in-person and virtual social interactions. We borrow William Little's definition of *social interaction*, which is the "process of reciprocal influence exercised by individuals over one another during social encounters." In face-to-face social interactions, humans utilize a set of verbal and non-verbal cues in order to initiate, respond to, and end interactions.

Various physical and interaction cues have evolved as common devices people utilize and operate across stages of social interactions. For example, eye contact and speech prosody can be used to

determine when people are about to start or yield their interaction during turn-taking [4, 6]. Eye gaze, being directly observable and accessible to others, can help people determine if social interactions should start or be averted by both conveying and taking in information from others [3]. Similarly, deictic gesturing in interactions can help to guide communication by directing and fostering joint attention [5]. Proximity can also be used to both portray and perceive familiarity, comfort, and attitude in social interactions [2].

During the evolution of our interconnectedness, virtual reality (VR) became an immersive space in which social interactions occurred, influencing and augmenting behaviors based on VR's affordances and limitations. While VR may be conventionally conceptualized as the technology that creates the effect of an interactive, audiovisual, and three-dimensional world [1], how social interactions may unfold in VR and what is its implication to the design of VR systems have not been thoroughly studied.

With the perceived presence of entities, including people and things, in an immersive space, people are offered with the opportunities to interact with one another in virtual environments. We conceptualize *Social VR* as socio-technical systems that allow people to participate in social encounters over immersive mediums in which objects, environments, and embodiments of other humans are co-present. Based on this working definition, examples of some currently available social VR platforms include VR Chat [1], AltSpace [2], RecRoom [3], and Mozilla Hubs [4], etc. in which people can enter into these virtual spaces with selected avatars using devices ranging from specialized VR headsets to generic web browsers. Resembling social interactions that take place in the physical world, social interactions in VR can occur with a combination of verbal and non-verbal cues, depending on what cues each platform makes available and how usable and accessible are these cues.

To identify interaction cues people may utilize in different stages of social interactions in VR, we conducted a pilot study that involved observing eight separate participants attending VR-based social events (e.g., virtual job fairs in one of the social VR platforms). Combined with common interactions cues known from existing literature, we identified eight potential cues related to the initiation and maintenance of social interactions: *changing head orientations*, *clapping hands*, *waving hands*, *moving closer to the person*, *greetings*

---

[1] https://hello.vrchat.com/
[2] https://altvr.com/
[3] https://recroom.com/
[4] https://hubs.mozilla.com/

(e.g., saying hello), ice breaker other than greetings (e.g., asking a random question to kick-start a conversation, for example, "excuse me, do you know how to..."), going into a topic directly, and *making some random noise to catch attention*; along with seven potential cues to end interactions: *changing head orientation, clapping hands, waving hands, moving away from the person, closing words (e.g., saying goodbye), other implicit cues of ending,* and *stop to respond.*

Given these potential interaction cues identified, we aim to explore the following research questions:

RQ1. What cues people use to initiate, respond to, and end a social interaction with others at VR-based events?

RQ2. Are there differences between the use of cues at VR-based events and those used at face-to-face events? If so, how do they compare?

2 METHOD

We collected data using an online survey to answer the above RQs. Detailed information about our sampling strategy and survey questions is provided below:

Sampling. We recruited participants from a public interest group of VR [5]. Qualified participate in our study include people who are a) 18 years old or above, b) currently based in the United States, and c) at one or more VR-based events over the past three months. We removed invalid responses that include questionable answers about one's VR experience (e.g., if a person claimed prior experience of attending VR-based events via a non-existing platform or on a future date). Our final pool of participants include 51 survey respondents (15 females, 36 males). Their average time to complete the survey was 13.8 min (SD = 7.1). Each participant received a $10 gift card for compensation.

Survey questions. Participants were asked to recall one VR-based event that they attended in the past three months, reflect on their experience at that event, and answer three groups of questions. These questions ask for the following aspects of information:

Overall goals and access to devices. Participants reported their goals of attending the VR-based event as well as the devices they used. These responses were collected via two multiple choice questions, respectively.

Use of cues to interact with others at VR-based events. Participants reported their extent of using various cue to initiate, respond to, and end interactions with others at VR-based event. They provide responses on 5-point Likert scales (1 = never, 5 = always).

Use of cues to interact with others at F2F events. Participants reported their extent of using various cue to initiate, respond to, and end interactions with others at F2F event. They provide responses on 5-point Likert scales (1 = never, 5 = always).

3 DATA ANALYSIS AND RESULTS

We describe our findings based on the quantitative analysis of survey responses.

3.1 Overall goals and access to devices

The majority of our participants (70.58%) attended VR-based events with one or multiple specific goals. Among them, the most common

---

[5]https://www.facebook.com/groups/Quest2Community/

goal is meeting new friends (64.70%), which is followed by experiencing VR (62.75%). Only 3.92% of the participants had no specific goal of attending the events. In terms of the access to devices, most of the participants (80.39%) reported that they joined the events using VR headsets, while the rest of them joined the events through computer browsers.

3.2 Use of interaction cues at VR-based events and F2F events

Cues to initiate a social interaction with others. We compared participants' use of various cues to initiate a social interaction at VR-based and F2F events, respectively (see Fig. 1 for full details). In VR-based events, participants reported *greetings* (e.g., saying hello) being the mostly used cue (M=3.61, SD=0.90) to initiate an interaction with others, while *moving closer to the person* was reported as being the mostly used one in F2F interactions (M=3.92, SD=0.74).

Further, we conducted paired sample t-tests to compare the use of cues when having social interactions in VR-based and F2F events. We found that participants used the cue of *moving closer to the person* more in F2F events (M=3.92, SD=0.74) than in VR-based events (M=3.22, SD=0.88), $t(50)=4.50$, $p<.01$. They also *made random noise to catch attention* more in F2F events (M=3.29, SD=1.17) than in VR-based events (M=3.00, SD=1.04), $t(50)=2.01$, $p<0.05$. In contrast, participants are more likely to *directly go into a topic* when initiating interactions in VR-based events (M=3.59, SD=1.02) than in F2F events (M=3.10, SD=0.90), $t(50)=2.57$, $p<0.05$. As for the usage of the other cues, no significant difference was found between VR and F2F events.

Cues to respond to others to maintain a social interaction. We compared participants' use of various cues when responding to others at VR-based events and F2F events, respectively (see Fig. 2 for full details). In VR-based events, participants reported *some sort of ice-breaker other than greetings* being the mostly used cue when responding to others (M=3.59, SD=0.96), while *moving closer to the person* was reported as being the mostly used one in F2F interactions (M=3.73, SD=1.02).

We conducted paired sample t-tests to compare the use of cues between the two conditions (i.e., VR-based vs. F2F events). We found that participants used the cue of *waving hands* more in VR-based events (M=3.45, SD=0.97) than in F2F events (M=3.12, SD=0.77), $t(50)=2.02$, $p<0.05$. They used *some sort of ice breaker other than greetings* more in VR-based events (M=3.59, SD=0.96) than in F2F events (M=3.25, SD=0.91), $t(50)=2.62$, $p<0.05$. They are also more likely to *directly go into a topic* when responding to others in VR-based events (M=3.53, SD=1.10) compared to in F2F events (M=3.20, SD=0.87), $t(50)=2.15$, $p<0.05$. As for the usage of the other cues in the two types of events, no significant difference was found.

Cue to end a social interactions with others. We compared participants' use of various cues to end a social interaction at VR-based events and F2F events, respectively (see Fig. 3 for full details). In VR-based events, participants reported *other implicit cues* (e.g., cues other than explicitly saying closing words such as "goodbye") being the mostly used one to end an interaction with others (M=3.41, SD=0.96). However, they mostly favored *explicitly say closing words* when ending interactions in F2F events (M=3.37, SD=0.94).
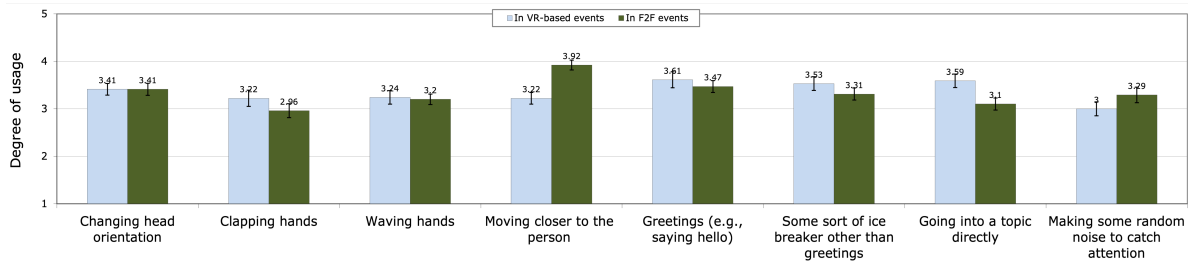
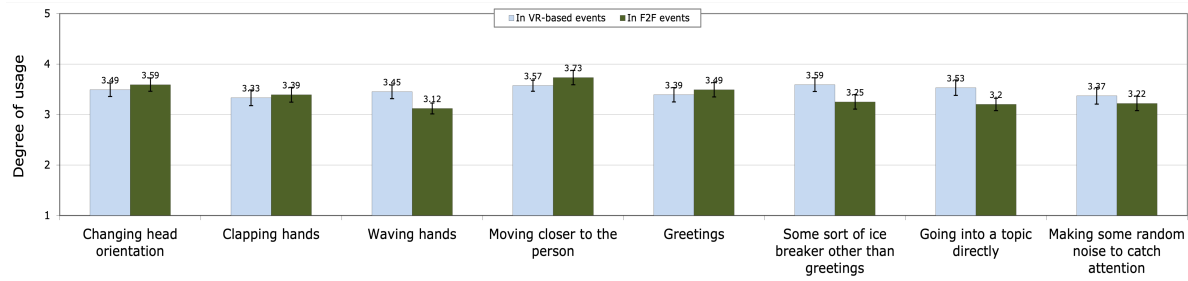**Figure 1: Cues used by participants when initiating social interactions in VR and F2F events.**

Figure 2: Cues used by participants when responding to social interactions in VR and F2F events.
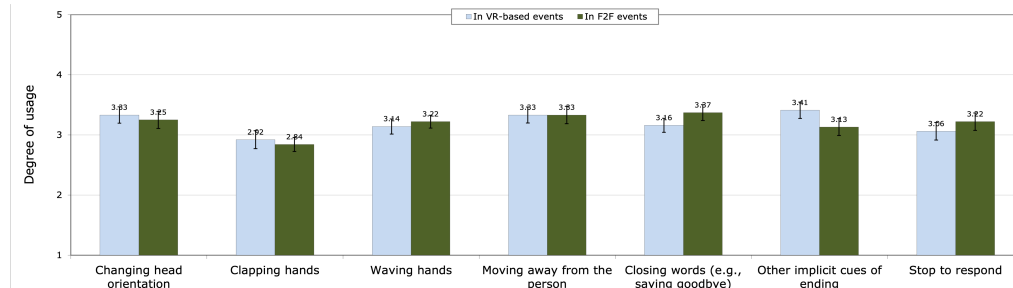

Figure 3: Cues used by participants when ending social interactions in VR and F2F events.

Paired sample t-tests did not reveal any significant differences between how participants use cues to end interactions in VR-based events and that in F2F events.

4   SUMMARY

In summary, our results suggest that people actively use diverse cues to facilitate their social interactions in VR-based social events, just as in face-to-face interactions. However, the degrees of using specific cues could vary depending on the medium (e.g., VR or F2F) used to undertake social interactions, and the stage of social interaction (e.g., initiating, responding or ending an interaction). For example, it's interesting to note that participants were more likely to directly go straight into a topic when initiating an interaction in VR, rather than using implicit non-verbal cues (e.g., moving closer to increase proximity or making sound to get attention) for initiations of interactions. Comparisons of cue usage in VR-based versus F2F social events further showed substantial differences on

how various cues were used during the initiation and responding stages between VR and F2F. While it is a commonly used strategy to use spatial proximity (moving closer to another person) and sound to foster or maintain social interactions in F2F encounters, these strategies were not as heavily used in VR. Instead, social interactions in VR appear to be more direct and straightforward than in F2F, potentially relying more on verbal strategies (e.g., using ice breakers) and simple non-verbal behaviors (e.g., waiving hands).

The results have initially answered our research questions concerning which cues people use most in VR and how cue usage in social VR resembles or differs from cue usage in F2F social interactions. However, what remains unclear are the motivations, mechanisms, norms, constraints, and opportunities people perceive and engage with when using various cues to initiate, maintain and end social interactions in virtual environments. Future work is needed to investigate these follow-up research questions to gain a

fuller picture of social interactions in VR and to shed light on the design of VR tools that better support online social events.

5 ACKNOWLEDGMENTS

REFERENCES

[1] S. Bryson. 2013. Virtual Reality: A Definition History - A Personal Essay. *ArXiv* abs/1312.4322 (2013), 6 pages.
[2] Gary W Evans and Richard E Wener. 2007. Crowding and personal space invasion on the train: Please don't make me sit in the middle. *Journal of Environmental Psychology* 27, 1 (2007), 90–94.
[3] Roy S Hessels. 2020. How does gaze to faces support face-to-face interaction? A review and perspective. *Psychonomic bulletin & review* 27 (2020), 856–881.
[4] Simon Ho, Tom Foulsham, and Alan Kingstone. 2015. Speaking and listening with the eyes: gaze signaling during dyadic interactions. *PloS one* 10, 8 (2015), e0136905.
[5] Gary M. Olson and Judith S. Olson. 2000. Distance Matters. *Hum.-Comput. Interact.* 15, 2 (Sept. 2000), 139–178.
[6] Maite Taboada. 2006. Spontaneous and non-spontaneous turn-taking. *Pragmatics* 16, 2-3 (2006), 329–360.
[7] Simon N Young. 2008. The neurobiology of human social behaviour: an important but neglected topic. *Journal of psychiatry & neuroscience: JPN* 33, 5 (2008), 391.