Exploring Human-to-Human Telepresence and the Use of Vibro-Tactile Commands to Guide Human Streamers

Kevin P. Pfeil(✉), Katelynn A. Kapalo, Seng Lee Koh, Pamela Wisniewski, and Joseph J. LaViola Jr.

University of Central Florida, Orlando, FL, USA
kevin.pfeil@knights.ucf.edu

Abstract. Human-to-human telepresence is rising to mainstream use, and there is opportunity to provide rich experiences through novel interactions. While previous systems are geared towards situations where two users are previously acquainted, or provide channels for verbal communication, our work focuses on situations where audio is not desirable or available, by incorporating vibro-tactile commands into a telepresence setup. We present results from a lab-based study regarding a human-to-human telepresence system which enables one person to remotely control another through these vibro-tactile cues. We conducted a study with 8 participants to solicit their feedback when acting as a Streamer, 8 additional participants to solicit feedback as a Viewer, and 30 bystanders, through surveys and debriefing sessions. While our participants generally found the application favorable, we did find mixed feelings towards vibro-tactile devices, and much room for improvement for the whole interaction. We discuss the implications of our findings and provide design guidelines for future telepresence developers.

Keywords: Telepresence · Vibro-tactile interface · 360° Video

1 Introduction

Telepresence is the ability to perceive and/or interact with a remote environment, as if actually there. Originally conceived by Minsky in 1980, it has been hypothesized that telecommunications can provide an avenue for a worker to perform a task from across the globe using a natural, egocentric point of view [31]. While this ultimate telepresence experience has not yet been fully realized, we have seen immense technological advances in that direction. For instance, the combination of 360° cameras and virtual reality (VR) head-mounted displays (HMDs) allow an individual to explore a remote environment with a greater sense of *immersion*, or a sense of being enveloped in an interaction (see [45,52]), than ever before. To facilitate mobile exploration, the robotics field has brought forth wheeled platforms which can be piloted through a simple graphical user

J. Y. C. Chen and G. Fragomeni (Eds.): HCII 2021, LNCS 12770, pp. 183–202, 2021.
https://doi.org/10.1007/978-3-030-77599-5_15

interface (GUI) [26]. However, these robots are typically constrained to a particular, pre-planned environment such as conference venues [36], hospitals [9], and sidewalks [12], as they still have problems in navigating difficult terrain.

To circumvent this difficulty, and to provide a more interpersonal experience, some researchers have hypothesized that the robot can be replaced with another human (called Streamer). This concept of human-to-human telepresence has recently emerged, and the HCI community is examining its benefits and detriments. However, we have still seen the implementation of human-to-human telepresence in the wake of the recent COVID-19 pandemic, which has forced worldwide communities to practice social distancing. For instance, the Faroe Islands conducted telepresence experiences by having a Streamer walk about the islands while remote users directed them where to go [15]. Additionally, some researchers suggest that we might even see this concept provide a new avenue for job creation; one person could perform tasks while being directed by a remote user. Misawa and Rekimoto describe a "physical body marketplace" [33], where the Streamer's role is equivalent to ride-sharing drivers (e.g. Uber and Lyft).

It is this style of interaction that thus influences our research interests. While numerous previous works have focused on the user experience for the Viewer (the person watching the video stream), we note that the user experience for the Streamer has been woefully understudied. One consideration is the Viewer may not have the ability to pick their partner, or perhaps they will share a Streamer's service with other Viewers simultaneously. As such, verbal guidance might not be desirable, or even possible. To help tackle this problem, we draw from Teleguidance literature, in which researchers evaluated vibro-tactile devices (in this paper, we use the shortened term VibTac). These devices use feedback in the form of vibrations to convey information (such as directional commands), and they have been shown to succeed for guidance and navigation as primary or auxiliary communication channels [8,19,28,30]. We suspect that they are suitable for use in human-to-human telepresence as well.

Therefore, in this paper, we explore this concept with a lab-based study with which we gather feedback from 16 participants regarding a telepresence setting using VibTac devices. We describe a prototypical interactive system that allows a Streamer to give a live-streamed tour to a Viewer, who in turn can provide navigational commands in the form of tactile cues through a VibTac belt or hat. Additionally, we surveyed 30 other observers to help understand their how the third-party perceives this style of interaction. Our primary research questions include the following:

- RQ1: What types of VibTac devices are preferred by Streamers?
- RQ2: How do third-party members perceive human-to-human telepresence interaction?
- RQ3: What scenarios would benefit from human-to-human telepresence?

We found that our participants did not have a particular preference between our VibTac devices, though the navigational commands were more strongly felt with our developed hat. Additionally, our participants responded that, although they were comfortable acting as a Streamer, they were not particularly energized to be one. Lastly, observers were somewhat comfortable being collocated with a Streamer, although they felt a greater sense of trust in the Streamer than they did in the remote Viewer.

2 Related Work

Telepresence Streamers have been leveraged in a variety of projects, but we have yet to understand how humans truly regard this role. In this section, we review the relevant literature at the intersection of Human-to-Human Telepresence and VibTac interfaces, and discuss our unique contributions to the telepresence community.

2.1 Human-to-Human Telepresence

Human-to-human Telepresence is a fairly new concept, as the technology and infrastructure to support this kind of interaction has just recently emerged; but, we have seen many projects that in part help to realize this concept. The JackIn system by Kasahara et al. shows how the use of omni-directional video cameras worn on the head by Streamers can give Viewers an immersive avenue to explore remote environments [16–18]. Here, the Viewer typically utilizes a VR HMD, granting the ability to explore the environment through natural head rotations. Pfeil et al. investigated how high a camera could be placed, and found that camera height is not a significant factor to consider for telepresence design [39,40]; thus, this interaction could be enjoyed by broad audiences. Even a simple live-streaming device, such as a mobile phone, can provide an adequate view, as with the Polly system which is mounted on a Streamer's shoulder [21,23,24]. Here, though the Viewer does not wear a VR HMD, they can change the viewpoint by controlling a gimbal that holds the camera, through a GUI. Between these two projects, the Viewer has either a first-person or third-person point of view in relation to the Streamer. The latter allows for a more personable connection between the two users, but this may or may not be desirable, as the Streamer could be a friend or family member, but they could be a stranger.

Some researchers have explored *who* can be a Streamer. Ishak et al. deployed a technology probe to the university classroom [14]. Here, the authors found that students would be interested in having their friends act as live Streamers, but there were reservations in having strangers be these proxies. This concern was in part addressed through works by Misawa and Rekimoto, where Streamers wore a tablet on their head, to display the face of the Viewer [32,33]. Here, the Streamer did not interact with the environment, except as instructed by the Viewer's audible instructions. In this way, there was an illusion that the Viewer was actually in the remote environment. The authors suggest that the optimal Streamer would be someone known to the Viewer, and someone who has similar physical traits, so to enhance the illusion.

One of the problems in the human-to-human telepresence literature is the common finding that the Streamers feel socially awkward when other people see them using this technology [20,38,42]. In all of the above examples, the devices are clearly identifiable by third-party users, resulting in this awkward feeling. It is thus important to find a way to balance user experience for both members. Baishya and Neustaedter envisioned a way for long distance couples to increase their togetherness using always-on video chat [3]. Here, a smartphone was placed in the shirt pocket of both users, with the camera facing out. In this way, it was inconspicuous to third party observers, although the opportunity to explore the remote environment is stymied because of where the camera is placed. As such, there is work that must be performed to reach an acceptable balance of user experience between all parties involved.

In our paper, we extend prior work by letting a Viewer command a Streamer through the use of VibTac devices. Our work is similar to that of Misawa and Rekimoto, in that the Streamer is asked to follow specific commands given by the Viewer [32,33], and our technology probe is not unlike that of the Faroe Islands remote tourism application [15]. However, our prototype devices were designed to be inconspicuous, by embedding them into clothing. Our work contributes to the telepresence literature by providing an understanding of Streamer's perception of human-to-human telepresence devices, and to understand how comfortable they are with the general interaction style.

2.2 Vibro-Tactile Interfaces

VibTac interfaces have been studied in the past, commonly in the form of belts [28,46,49,50]; but there have also been others in the form of vests [29], helmets [35], head bands [6], caps [19], and even socks and shoes [30]. They have been integrated into user interfaces to support the visually impaired; for instance, McDaniel et al. described their belt prototype in an effort to convey where other people were, in both direction and proximity [28], and Wang et al. developed a belt that provided directional commands to help users navigate around obstacles without a cane [51]. In these implementations, computer vision techniques are used to identify detect people and obstacles, and to calculate a route around them.

However, algorithms do not operate at 100% accuracy. It is sometimes recommended to have a human-in-the-loop (HITL) interaction, so to leverage human expertise and correct problems when algorithms make mistakes. Scheggi et al. proposed an HITL system in which a remote expert could view the surroundings of a visually impaired person, and provide real-time feedback in the form of vibrations [43]. Our work is inspired by this prior work, to help understand if VibTac devices can be used for general telepresence use. There is not a wide range of consumer-grade VibTac wearables, so we developed two simple devices; one is a belt similar to Tsukada's ActiveBelt [49], and the other is a hat based on a reduction of Kaul's HapticHead device [19]. Literature does not seem to point to optimal tactor configurations, so iterative development helped direct our design of these devices. In this paper, we elicit Streamer feedback to the VibTac belt and hat, to understand which provides the clearest set of instructions.

3 Prototype System Design

We developed a custom live-streaming application in Unity3D; Unity's multiplayer service allowed us to easily transmit data over the network. To provide visuals, we used the Ricoh Theta S 360° camera, which supports live streaming. The camera was mounted to a backpack rig, giving an overhead view of the Streamer, similar to the work by Tang et al. [48]. We modified FFPLAY[1] code and used it as a DLL in Unity. A script decoded the live, equirectangluar video frames, turned them into RGB values, and mapped them into a spherical image. We were able to achieve a streaming rate of 30 frames per second with approximately 1 s of latency. Using a virtual reality head-mounted display (HMD), or by simply click-and-dragging with a mouse, a Viewer could manipulate the viewpoint to any desired angle. See Fig. 1 for a visual of the hardware worn by the Streamer; this consisted of a backpack rig, the pole-mounted camera, a VibTac device, and a laptop to send and receive data over the network. The total weight was approximately 20lbs (9.1 kg).

The VibTac devices we developed were a belt and a hat (see left section of Fig. 1). Each utilized four coin vibration motors that were sewn into the clothing and interfaced with an Arduino Uno, which was programmed to read serial port data. The Uno was plugged into the Streamer's computer via USB, and Unity's multiplayer service allowed us to easily transmit commands. When activating the motors, we used 3.3V (output of the digital pins), as we found 5V to burn the motors. The motors did not make direct contact with the body, and were veiled behind a layer of fabric. This is in accordance with the findings by Kaul and Rohs, that direct contact could cause marks on the skin of the user [19]. Our navigational commands were similar to those used by the Faroe Islands remote tourism [15]; we simply needed our devices to convey four commands - *Go Forward*, *Go Backward*, *Turn Left*, and *Turn Right*.

For the belt, the mounted motors were sewn into an elastic band that used a belt buckle for fastening. Our final design called for motors contacting the belly, the back, and each hip. Activating the belly motor indicated *Go Forward*, activating the right hip motor meant *Turn Right*, and so forth. For the hat, the motors were sewn into the fabric - our final design consisted of two placed near the temples, and two placed near the sides of the neck. Activating the front motors indicated *Go Forward*, activating the two right motors meant *Turn Right*, and so forth.

When a Viewer pressed a navigation button, a command was sent to the Streamer's laptop, activating the appropriate motors. Latency from button press to motor activation was less than 1 s. For our study, we utilized two Viewer conditions. The first condition used the HTC Vive VR headset, where the built-in head tracking allowed the user to change viewpoints while watching the 360° camera feed. The second condition used a flat laptop screen, and by click-and-dragging a mouse pointer, the omnidirectional video feed could be manipulated in real-time.

[1] https://ffmpeg.org.

Fig. 1. **Left**: The developed tactile belt (top) and hat (bottom). Each utilized a 4-tactor design in order to provide feedback to the Streamer. The motors were controlled by an Arduino Uno. **Right**: A user wearing the backpack rig, which consisted of a 360 camera for livestreaming, a vibro-tactile device for feedback, and a laptop.

4 User Study

We conducted a lab-based study at the University of Central Florida with IRB approval. The study aimed at garnering user feedback from Streamers, Viewers, and other third-party members who were not part of the interaction.

4.1 Study Design

We conducted a 2×2 mixed-design study; participants were split into 2 groups - one group consisted of the Streamers, in which the VibTac device selection was a within-subjects variable. The other group consisted of Viewers, in which the Viewing Mode was a within-subjects variable. Per group, condition order was counter-balanced. In addition, we conducted a short survey with non-participants. Our dependent variables aimed at measuring the perceived value of this kind of human-to-human telepresence, effectiveness of VibTac devices as a primary communication channel, and perception of the Viewer-Streamer relationship.

4.2 Subjects and Apparatus

We recruited eight participants to act as a Viewer. These participants consisted of 4 males and 4 females, and their age ranged between 18 and 26 (M = 21.8; SD = 3.20). On a 5-point Likert scale where 1 means Never and 5 means Always, the participants very rarely watched 360° videos (M = 1.75; SD = 0.46). They again very rarely used VR (M = 1.89; SD = 0.64). They very rarely drove robots (M = 1.89; SD = 0.64), and they rarely played first-person video games (M = 2.25; SD = 1.04).

We recruited eight participants to act as a Streamer. These participants consisted of 6 males and 2 females, and their age ranged between 18 and 38 (M = 24.4; SD = 6.30). Only 1 of our participants had experience with a wearable device (such as an Apple Watch or Fitbit). On a 5-point Likert scale where 1 means Never and 5 means Always, the participants exercised somewhat often (M = 3.5; SD = 0.76). They sometimes felt phantom vibrations regarding their cell phone (M = 2.88, SD = 0.83).

We surveyed an additional 30 university students (17 female) to understand their thoughts regarding the concept of human-to-human telepresence. All of our participants were recruited from the University of Central Florida.

4.3 Procedure

Viewers and Streamers. Our recruited participants reported to our lab, where they reviewed an informed consent form. After agreeing to participate, they were given a demographics survey, followed by a brief description of the telepresence setup. We explained how the VibTac interfaces worked, and how the controls were akin to a computer game (WASD keys). After introducing the hardware, our participants moved to their assigned position.

At any given time, a participant was accompanied by an experimenter, and we did not have interaction between participants. During Streamer trials, an experimenter provided the navigational commands. During Viewer trials, an experimenter served as the Streamer. As such, each participant only interacted with authors.

The Streamer was guided through a building on our university campus. The building layout consists of multiple hallways and corridors, which allowed us to take advantage of all VibTac commands. When controlling a Streamer participant, we held down the commands such that the vibration effect was constant. The *Stop Moving* command was issued when no vibrations were activated. We did this to ensure our Streamers had ample interaction time with the VibTac interfaces.

The navigational route was randomized and kept secret from the Streamer. As such, they had to rely on the VibTac devices to complete the trial; Streamers were asked to walk at a comfortable pace, and not run. After a Streamer completed a route, they changed VibTac devices and was then navigated through a different route. After a Viewer participant completed a route, they changed viewing modes and then performed navigation through a different route. In total, a Streamer participant was navigated twice; similarly, a Viewer participant conducted navigation twice. After each condition, a questionnaire was administered to elicit feedback regarding that particular condition. After both trials were complete, a final questionnaire was administered to garner feedback in terms of preference. Between participants, we cleaned the HMD and hat instruments with rubbing alcohol to ensure sanitation. The study took approximately 30 min to complete, and participants were paid $5 for their time. This study was conducted prior to the COVID-19 pandemic.

Survey Respondents. Survey respondents were approached in public and asked if they would like to provide their comments regarding human-to-human telepresence. One author wore the backpack rig, and we described the concept. Feedback was received in the form of a questionnaire and verbal communication. Respondents were not compensated.

4.4 Soliciting User Feedback

We administered quantitative measures to help understand user feedback, in terms of enjoyment and comfort, as well as social and economic aspects; see Table 1 for a list of questions asked to our participants and respondents. All closed-ended questions were given on a 7-pt Likert scale where 1 meant "Strongly Disagree" and 7 meant "Strongly Agree". Open-ended questions were given as free-response prompts.

For the Viewer participants, we also asked simple questions to see if there was a preference in Viewing Mode, but we were more interested in their thoughts regarding the concept of controlling another human. For our Streamer participants, we were also interested in understanding VibTac device preference. We asked questions to understand how well the devices conveyed instructions and how comfortable they were.

4.5 Data Analysis Approach

As we did not have enough power to run statistical tests within the Streamer and Viewer groups, we report our results using descriptive statistics. To determine the overall effectiveness of each VibTac device, we averaged the 4 directional command responses into an index, per device. For the responses garnered from the final questionnaire, we report descriptive statistics and also discuss the positive/negative aspects as indicated by our users.

Table 1. Questions administered to participants. The Group column indicates which set of participants received the question: V = Viewers; S = Streamers; R = Survey Respondents. Asterisks (*) denote questions were given for multiple conditions.

Question	Group
I felt comfortable knowing I was controlling another human	V
I liked having control over the Streamer	V
I trusted the Streamer to go where I wanted them to go	V
I would feel comfortable navigating the Streamer into a private place, such as a restroom	V
I felt comfortable knowing I was being controlled by another human	S
I liked being controlled by the Viewer	S
I trusted the Viewer to give me good directions	S
I would feel comfortable being navigated into a private place, such as a restroom	S
It was easy to understand the Viewer's directions	*S
I could easily feel the *Move Forward* command	*S
I could easily feel the *Move Backward* command	*S
I could easily feel the *Turn Left* command	*S
I could easily feel the *Turn Right* command	*S
Assuming the Streamer did a good job, I would consider paying for a live virtual tour	V, R
How many dollars (USD) would you pay for a live virtual tour?	V, R
Assuming I would be adequately compensated, I would consider being a Streamer as a job	S, R
How many dollars (USD) would you expect to receive, to be a Streamer?	S, R
I would trust the Streamer to do the right thing	R
I would trust the Viewer to do the right thing	R
I think it would be fun to be a Streamer	R
I think it would be fun to be a Viewer	R
I would be comfortable knowing a Streamer is near me in a public place	R
I would be comfortable knowing a Streamer is near me in a private place	R
What did you like about the system?	V, S
What did you dislike about the system?	V, S
What changes would you make to improve the overall experience?	V, S
What scenarios would you use this system for?	V, S, R

5 Results

In this section we report the results of our study, split by the different roles. We begin with the Viewer participants, followed by the Streamer participants, and finish with the survey respondents.

5.1 Viewer Considerations

Overall, our Viewers liked using the system and they had no major discomforts with either Viewing Mode. Six participants were most comfortable with the flat screen, while only two were more comfortable with the HMD. However, five indicated that they preferred the HMD to the flat screen.

Our Viewer participants felt comfortable controlling the Streamer (M = 5.5, SD = 2.1) and typically liked the interaction (M = 5.4, SD = 2.1). Only one participant responded negatively, as they were afraid that they would cause harm to the Streamer: *"Controlling people scared me half to death! The building is new to me and I was extremely scared I was going to run the Streamer off a ledge"*. This participant was the only one to express consideration for the Streamer.

Generally, our participants had great trust that the Streamer would follow the commands accurately (M = 6.6, SD = 0.5). However, the participants were very apprehensive about the idea of sending a Streamer into a private place; see Fig. 4, with only one participant feeling comfortable regarding that situation. Considering the concept as a whole, participants responded that they would be likely to hire a Streamer for a tour (M = 5.5, SD = 1.6). The range of value was very broad, however; the smallest sum of money participants were willing to pay was $5, and the most was $25 (see Fig. 3).

Our open-ended questions revealed that the most negative features of our telepresence setup was the latency; half of our participants complained about lag. From button press to motor activation, approximately 1 s elapsed; but, there was additional overhead time from motor activation to Streamer response, followed by another second of video lag. As such, there were many times where

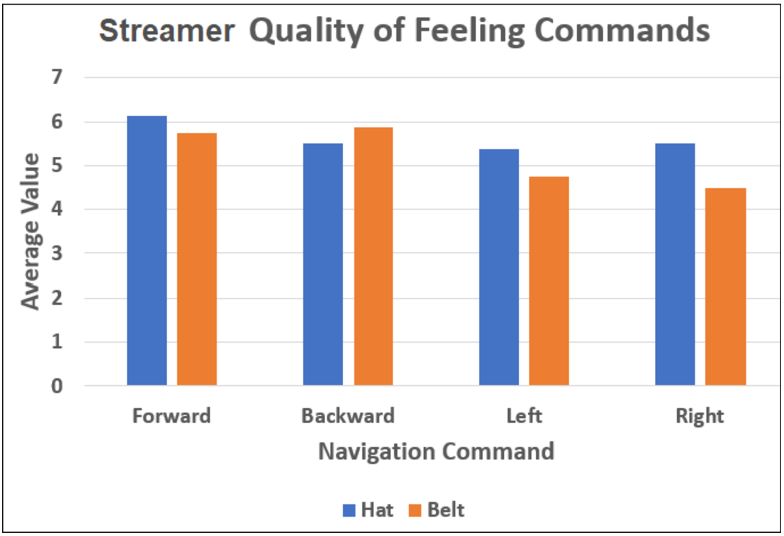

Fig. 2. The commands from both VibTac devices were perceived well, but the Hat device was slightly more conducive to feeling the vibrations.

the Streamer walked too far down a hallway, or missed a turn. In addition to better visuals, some of our participants indicated that they would prefer to hear and speak to the Streamer. Usage scenarios for human-to-human telepresence included campus orientation for new or prospective students, playing augmented reality games, virtual tourism, and remote shopping.

5.2 Streamer Considerations

Our participants generally had no issue with either VibTac device. Comfort was comparable between the Belt (M = 5.4, SD = 1.3) and the Hat (M = 5.4, SD = 1.2); likewise, users felt only slightly embarrassed when wearing either device (Belt: M = 2.3, SD = 1.5; Hat: M = 2.6, SD = 1.8). Regarding the quality of VibTac commands, there was only a slight difference in the ability to detect them (Belt: M = 5.2, SD = 1.5; Hat: M = 5.6, SD = 1.0), and regardless of device, participants felt that they understood where they were being directed to go (Belt: M = 5.3, SD = 1.7; Hat: M = 5.8, SD = 1.0). Our participants indicated that they generally felt the belt vibrations with ease, but the hat vibrations were more distinct (see Fig. 2). As such, device preference was not in favor of one over the other; three users preferred the belt, and five preferred the hat.

Participants indicated that they enjoyed the novelty of our system, and we suspect that this novelty may have impacted participant Likert-scale responses. However, open-ended questions did help to reveal weaknesses of our devices. Five of our participants indicated that while the vibrations were enough to complete the walking task, they could have been better. For example, the commands from the belt were perceived as having different strengths. One user informed us that they needed to put their hands on the belt in order to recognize the vibrations. On the other hand, while perception towards the hat was comparable to the belt, users indicated that the vibrations caused a tickling sensation, or were more distracting. This is due to the motors being close to the users' forehead and ears, so the vibrations were more distinct and audible. One user suggested that they were able to hear the motors vibrating near their ears, but this helped them to determine which direction was requested.

5.3 Survey Respondents

Our survey respondents, who were third-party bystanders, had mixed feelings regarding this concept. While they had general trust in a Streamer (M = 5.3, SD = 1.1), they were much less trusting in the Viewer (M = 4.4, SD = 1.7). A Wilcoxon Signed Rank Test revealed a significant difference in who they would trust ($Z = -2.964, p < .005$). This is because they are able to see and perhaps identify the local user, but they have no indication as to who is on the other end of the interaction. As such, they were somewhat comfortable with the idea of a Streamer being nearby in a public place (M = 4.9, SD = 1.3), but were far less enthused with the idea of a Streamer being collocated in a private area (M = 3.3, SD = 1.6). A Wilcoxon Signed Rank Test reveals a significant difference in the comfort of these scenarios ($Z = -4.168, p < .001$). Although they did

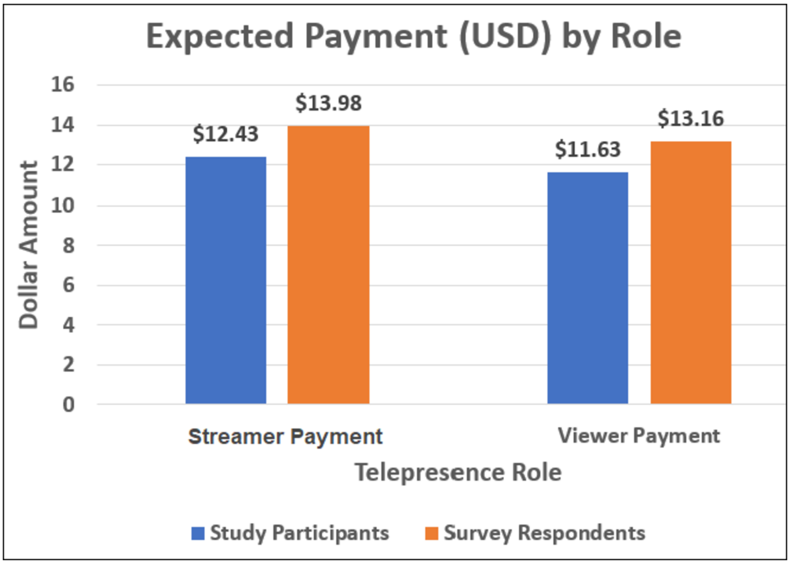

Fig. 3. Expected payment (USD) per telepresence role. Participants and survey respondents both indicated that the Streamer would garner more compensation than Viewers should pay.

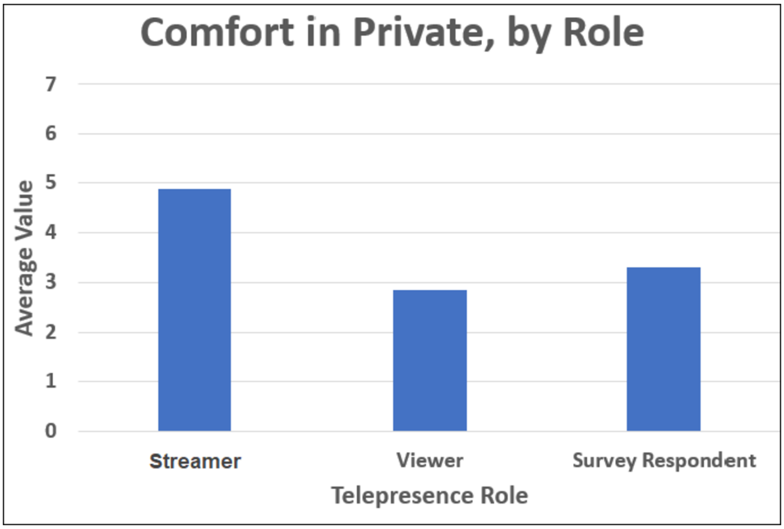

Fig. 4. Average response to comfort of telepresence in a private setting. Interestingly, the Streamers had little issue with the concept, whereas the Viewers and survey respondents were much less enthused.

not see a telepresence scenario taking place, the respondents thought it would be somewhat fun to be a viewer (M = 5.8, SD = 1.4), and less fun to be a Streamer (M = 5.2, SD = 1.6). A Wilcoxon Signed Rank Test revealed a significant difference between these sentiments ($Z = -2.056, p < .05$). Interestingly, however, they responded with slight interest in being a Streamer for payment (M = 5.0, SD = 1.8), and were less interested in paying for a live tour (M = 4.4, SD = 1.6). A Wilcoxon Signed Rank Test revealed a significant difference of interest for these roles ($Z = -2.024, p < .05$). The range of expected dollars-per-hour payment for being a Streamer was broad; the smallest rate was $7, and the largest was $30 (M = 13.98, SD = 5.80). The dollars-per-hour range for being a viewer was from $4 to $20 (M = 13.15, SD = 7.15). Thus, the respondents felt that a Streamer would warrant more compensation than a viewer would need to pay. See Fig. 3 for illustration. The respondents had many scenarios in mind for human-to-human telepresence, including tele-health physical therapy, remote exploration of live events like concerts, housing and campus tours, vlogging, tourism, military missions where talking is prohibited, and giving directions.

6 Discussion

In this section, we unpack the findings which emerged from our study, to address our research questions. We begin with how all parties view each other; next we discuss the perceived value of the interaction; third we describe the potential scenarios pondered by our users; and last we discuss VibTac device efficiency.

6.1 RQ1: Towards Better Telepresence VibTac Devices

Although our Streamer participants responded (with bias) in favor of the VibTac devices, we learned from them that there is room for improvement. First, there are benefits and drawbacks for each device. The belt was not as conducive as the hat in terms of interpreting the commands. McDaniel et al. found users to easily localize the vibrations emitted from their 7-tactor belt [28], but our participants had a measure of uncertainty in simply feeling the vibrations. Although ours and their tactors each vibrated on 3V, theirs were felt more strongly. While we cannot confirm, we believe that this difference stems from the distinct procedures between our studies. While their participants were standing still, ours were engaged in a walking task. As such, our participants had more distractions as they navigated their route. To overcome this problem, we plan on reiterating our belt with the use of motors which can withstand greater voltage ratings.

The hat allowed our participants to recognize the commands more strongly than the belt, but it was not perceived to be perfect. First, the constant vibration of the motors against the cranium caused a tickling sensation to some, a result also found by Kaul and Rohs with HapticHead [19]. They suggest to decrease the frequency in which the motors activate, and we believe that would have caused more comfort with our users. We purposefully kept the vibrations constant during interaction to elicit user responses, but in a next iteration we would transition hat vibrations towards more gestural commands.

Although neither of our VibTac devices were "perfect" according to our participants, they were still able to convey commands properly and discretely. We do not anticipate for one VibTac wearable to emerge as an optimal device, and believe that it is more appropriate to provide a range of options for a Streamer to use.

6.2 RQ2: Third-Party More Trusting of Streamer Than of Viewer

Our survey respondents were interestingly more trusting of a Streamer than of a remote Viewer. Although the Streamer will be the individual who is wearing the hardware, respondents see remote Viewers as the potential wrong-doers; they are the ones consuming the video stream and potentially making recordings. This is exemplified with the general discomfort of the idea of being collocated with a Streamer in private. Still, respondents were not particularly thrilled with the idea of being collocated with a Streamer in public. Prior research has found a negative attitude towards live streaming, especially without consent [27]; but in countries where public video recording is protected, e.g. in the United States, there is no real obligation to inform people that they might be in a stream [25, 47]. With technological advances, some researchers have begun asking questions regarding streaming ethics [10]; future research should target human-to-human telepresence to help understand how Streamers and their remote Viewers are perceived in the scope of ethics and legality.

6.3 Imbalance of Streamer/Viewer Payment

While monetary exchange is not necessary for friends and family to use telepresence, it is implied that it would help to bring about more general use. Our survey respondents indicated an imbalance regarding monetary payment and compensation between the viewer and Streamer, in that they would expect a Streamer to receive more money than they would be willing to pay. This is an issue that would prevent a "physical body marketplace", as conceived by Misawa and Rekimoto [33], from becoming reality, at least for interactions that involve 1 Streamer and 1 viewer. If this is to become a mainstream interaction style, it is clear that there needs to be more incentive to participate, in both ends, as suggested by prior work [5,11]. Future research can help to identify avenues which will provide these incentives; as an example, some researchers are focusing on ways to increase immersion through multisensory stimulation, including touch [22,37] and smell [4,34], which could lead to a more enjoyable experience and thus a stronger desire to participate. Additionally, it may be the case where a 1-to-1 interaction might not be suitable for human-to-human telepresence. Instead, by adopting a 1-to-many paradigm, where a single Streamer could give tours to multiple viewers simultaneously, the cost per viewer could decrease. Although there are additional challenges to be met in this type of interaction, it is possible to achieve. For example, the Faroe Islands remote tourism experience gives control to a single viewer at a time [15], but as such, control time is limited. We look forward to future studies and design ideas regarding how to improve upon the Streamer-viewer experience, to make it desirable by all parties.

6.4 RQ3: Expected Interaction Scenarios

Our participants and survey respondents were able to conceptualize a broad range of scenarios in which human-to-human telepresence may prosper, including more personal cases such as physical therapy or playing augmented reality games with a friend as a Streamer. More intimate scenarios such as giving directions and physical therapy can be found in prior research [7,41], and popular social media sites such as FaceBook, YouTube, and Twitch.tv provide platforms for larger audience engagement such as vlogging of travels and activities [1]. Popular live streams (especially those for games, found on Twitch) typically have a specific goal or direction, but some streamers do poll their viewers to provide more personalized content. Viewers watch these streams for multiple reasons [13,44], including to live vicariously through the experiences of another person [2], but with human-to-human telepresence, Viewers have the opportunity to engage in an even more personal experience. We suggest that the creation of a platform specific to one-on-one telepresence would help create more engaging experiences which are currently unavailable.

6.5 VibTac Devices Not Optimal for Primary Communication Mode

The novelty of VibTac-based navigation gave rise to positive feedback regarding our belt and hat. We did hear some suggestions for improving the devices, including a way to make them more inconspicuous (e.g. replacing the wires with a Bluetooth connection); but, our short-term study did reveal some disinterest with solely relying on VibTac as a primary mode of navigation. Participants on both sides of the interaction expressed a desire for audible communication with their interaction partner. Additionally, our users were on the cusp of becoming annoyed with constant vibrations. As such, we would recommend telepresence designers to consider adding VibTac as an auxiliary mode of communication, as well as exploring additional modes not studied here. As wearable VibTac devices have been broadly researched with positive findings [19,28–30,35], there is opportunity to let Streamer users pick their own as an option.

7 Limitations and Future Work

Although our interaction prototype was met with positive feedback, we did also find apprehensions. As such, there is still much work to be done on many fronts. Our results are from a test where the participants and authors met before conducting the experiment. We do not assume our results to generalize to situations where both parties are absolute strangers. As such, field tests in real-world scenarios are needed to further our understanding; but, ethical considerations must be made to ensure the rights of all parties (viewers, streamers, and third-party) are not infringed upon. Regarding Streamer control, it remains to be seen how users react to extended lengths of interaction exposure. Longitudinal studies should also be conducted to help identify problems which arise over time. Further, our study did not utilize pre-validated instruments, and instead offer insight through custom measures. In our future work, we will identify proper instruments to measure perception towards VibTac devices.

Our prototype was relatively bulky, and iterative ergonomic enhancements can be made to ensure Streamer comfort. Additionally, although we chose to study two of the most prevalent haptic devices found in previous literature, there are others which need to be thoroughly studied. Lastly, legal and ethical considerations must be investigated. In spirit, our study assumed that the interaction would be performed as intended. However, it is possible for a Streamer to be directed to questionable locations, or perform questionable actions (such as commit a crime). Though we do not want to instigate such a scenario, the telepresence community must ask what the proper response would be. We hope that our paper contributes to the discussion for this type of interaction which is rapidly approaching the mainstream.

8 Conclusion

We have presented our prototype in which a person can direct another through vibro-tactile commands. Our results indicate that there is a real opportunity to provide novel and desirable interaction, but more work is needed to make that a widespread reality. We envision this type of system becoming a prominent avenue in social media, allowing Viewers to have a more direct form of experience through the use of a Streamer. We look forward to seeing how this type of technology will engage users of all backgrounds, in order to explore the world around them.

Acknowledgments. Special thanks to Ravikiran Kattoju for assistance with running the user study.

References

1. Alohali, A., Kunze, K., Earle, R.: Run with me: designing storytelling tools for runners. In: Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct, pp. 5–8. ACM (2016)
2. Anderson, K.E.: Getting acquainted with social networks and apps: streaming video games on twitch. tv. Library Hi Tech News **35**(9), 7–10 (2018)
3. Baishya, U., Neustaedter, C.: In your eyes: anytime, anywhere video and audio streaming for couples. In: Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing, pp. 84–97. ACM (2017)
4. Baus, O., Bouchard, S.: Exposure to an unpleasant odour increases the sense of presence in virtual reality. Virtual Reality **21**(2), 59–74 (2017)

5. Bellotti, V., Ambard, A., Turner, D., Gossmann, C., Demkova, K., Carroll, J.M.: A muddle of models of motivation for using peer-to-peer economy systems. In: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, pp. 1085–1094. ACM (2015)
6. Cassinelli, A., Reynolds, C., Ishikawa, M.: Augmenting spatial awareness with haptic radar. In: 2006 10th IEEE International Symposium on Wearable Computers, pp. 61–64. IEEE (2006)
7. Chaudary, B., Paajala, I., Keino, E., Pulli, P.: Tele-guidance based navigation system for the visually impaired and blind persons. In: Giokas, K., Bokor, L., Hopfgartner, F. (eds.) eHealth 360°. LNICST, vol. 181, pp. 9–16. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-49655-9_2
8. Cosgun, A., Sisbot, E.A., Christensen, H.I.: Guidance for human navigation using a vibro-tactile belt interface and robot-like motion planning. In: 2014 IEEE International Conference on Robotics and Automation (ICRA), pp. 6350–6355. IEEE (2014)
9. Ellison, L.M., Nguyen, M., Fabrizio, M.D., Soh, A., Permpongkosol, S., Kavoussi, L.R.: Postoperative robotic telerounding: a multicenter randomized assessment of patient outcomes and satisfaction. Arch. Surg. **142**(12), 1177–1181 (2007). https://doi.org/10.1001/archsurg.142.12.1177
10. Faklaris, C., Cafaro, F., Hook, S.A., Blevins, A., O'Haver, M., Singhal, N.: Legal and ethical implications of mobile live-streaming video apps. In: Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct, pp. 722–729. ACM (2016)
11. Glöss, M., McGregor, M., Brown, B.: Designing for labour: uber and the on-demand mobile workforce. In: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems, pp. 1632–1643. ACM (2016)
12. Heshmat, Y., et al.: Geocaching with a beam: shared outdoor activities through a telepresence robot with 360 degree viewing. In: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, p. 359. ACM (2018)
13. Hilvert-Bruce, Z., Neill, J.T., Sjöblom, M., Hamari, J.: Social motivations of live-streaming viewer engagement on twitch. Comput. Hum. Behav. **84**, 58–67 (2018)
14. Ishak, C., Neustaedter, C., Hawkins, D., Procyk, J., Massimi, M.: Human proxies for remote university classroom attendance. In: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems, pp. 931–943. ACM (2016)
15. Islands, V.F.: Remote tourism - visit faroe islands (2020). https://visitfaroeislands.com/remote-tourism/
16. Kasahara, S., Nagai, S., Rekimoto, J.: Jackin head: immersive visual telepresence system with omnidirectional wearable camera. IEEE Trans. Visual Comput. Graphics **23**(3), 1222–1234 (2017)
17. Kasahara, S., Rekimoto, J.: Jackin: integrating first-person view with out-of-body vision generation for human-human augmentation. In: Proceedings of the 5th Augmented Human International Conference, p. 46. ACM (2014)
18. Kasahara, S., Rekimoto, J.: Jackin head: immersive visual telepresence system with omnidirectional wearable camera for remote collaboration. In: Proceedings of the 21st ACM Symposium on Virtual Reality Software and Technology, pp. 217–225. ACM (2015)
19. Kaul, O.B., Rohs, M.: Haptichead: a spherical vibrotactile grid around the head for 3D guidance in virtual and augmented reality. In: Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, pp. 3729–3740. ACM (2017)

20. Kim, S., Junuzovic, S., Inkpen, K.: The nomad and the couch potato: enriching mobile shared experiences with contextual information. In: Proceedings of the 18th International Conference on Supporting Group Work - GROUP 2014, Sanibel Island, Florida, USA, pp. 167–177. ACM Press (2014). http://dl.acm.org/citation.cfm?doid=2660398.2660409. https://doi.org/10.1145/2660398.2660409
21. Kimber, D., et al.: Polly: telepresence from a guide's shoulder. In: Agapito, L., Bronstein, M.M., Rother, C. (eds.) ECCV 2014. LNCS, vol. 8927, pp. 509–523. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-16199-0_36
22. Kontaris, D., Harrison, D., Patsoule, E.E., Zhuang, S., Slade, A.: Feelybean: communicating touch over distance. In: CHI 2012 Extended Abstracts on Human Factors in Computing Systems, pp. 1273–1278. ACM (2012)
23. Kratz, S., Avrahami, D., Kimber, D., Vaughan, J., Proppe, P., Severns, D.: Polly wanna show you: examining viewpoint-conveyance techniques for a shoulder-worn telepresence system. In: Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct, pp. 567–575. ACM (2015)
24. Kratz, S., Kimber, D., Su, W., Gordon, G., Severns, D.: Polly: being there through the parrot and a guide. In: Proceedings of the 16th International Conference on Human-Computer Interaction with Mobile Devices & Services, pp. 625–630. ACM (2014)
25. Kreimer, S.F.: Pervasive image capture and the first amendment: memory, discourse, and the right to record. U. Pa. L. Rev. **159**, 335 (2010)
26. Kristoffersson, A., Coradeschi, S., Loutfi, A.: A review of mobile robotic telepresence. In: Advances in Human-Computer Interaction 2013, p. 3 (2013)
27. Li, Y., Kou, Y., Lee, J.S., Kobsa, A.: Tell me before you stream me: managing information disclosure in video game live streaming. In: Proceedings of the ACM on Human-Computer Interaction 2(CSCW), no. 107 (2018)
28. McDaniel, T., Krishna, S., Balasubramanian, V., Colbry, D., Panchanathan, S.: Using a haptic belt to convey non-verbal communication cues during social interactions to individuals who are blind. In: IEEE International Workshop on Haptic Audio visual Environments and Games, HAVE 2008, pp. 13–18. IEEE (2008)
29. McGrath, B., Estrada, A., Braithwaite, M., Raj, A., Rupert, A.: Tactile situation awareness system flight demonstration. Technical report, Army Aeromedical Research Lab Fort Rucker AL (2004)
30. Meier, A., Matthies, D.J., Urban, B., Wettach, R.: Exploring vibrotactile feedback on the body and foot for the purpose of pedestrian navigation. In: Proceedings of the 2nd International Workshop on Sensor-Based Activity Recognition and Interaction, p. 11. ACM (2015)
31. Minsky, M.: Telepresence (1980)
32. Misawa, K., Rekimoto, J.: Chameleonmask: embodied physical and social telepresence using human surrogates. In: Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems, pp. 401–411. ACM (2015)
33. Misawa, K., Rekimoto, J.: Wearing another's personality: a human-surrogate system with a telepresence face. In: Proceedings of the 2015 ACM International Symposium on Wearable Computers, pp. 125–132. ACM (2015)
34. Munyan, B.G., Neer, S.M., Beidel, D.C., Jentsch, F.: Olfactory stimuli increase presence during simulated exposure. In: Lackey, S., Shumaker, R. (eds.) VAMR 2016. LNCS, vol. 9740, pp. 164–172. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-39907-2_16

35. Myles, K., Kalb, J.T.: Guidelines for head tactile communication. Technical report, Army Research Lab Aberdeen Proving Ground Md Human Research and Engineering Directorate (2010)
36. Neustaedter, C., Venolia, G., Procyk, J., Hawkins, D.: To beam or not to beam: a study of remote telepresence attendance at an academic conference. In: Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing, pp. 418–431. ACM (2016)
37. Pallarino, T., Free, A., Mutuc, K., Yarosh, S.: Feeling distance: an investigation of mediated social touch prototypes. In: Proceedings of the 19th ACM Conference on Computer Supported Cooperative Work and Social Computing Companion, pp. 361–364. ACM (2016)
38. Pfeil, K., Chatlani, N., Wisniewski, P.: Bridging the socio-technical gaps in body-worn interpersonal live-streaming telepresence through a critical review of the literature. Proc. ACM Hum.-Comput. Interact. (2021, to appear)
39. Pfeil, K., Wisniewski, P., LaViola Jr., J.J.: An analysis of user perception regarding body-worn 360° camera placements and heights for telepresence. In: ACM Symposium on Applied Perception 2019. SAP 2019. Association for Computing Machinery, New York (2019). https://doi.org/10.1145/3343036.3343120
40. Pfeil, K., Wisniewski, P.J., Laviola Jr., J.J.: The effects of gender and the presence of third-party humans on telepresence camera height preferences. In: ACM Symposium on Applied Perception 2020. SAP 2020. Association for Computing Machinery, New York (2020). https://doi.org/10.1145/3385955.3407924
41. Pulli, P., et al.: Mobile augmented teleguidance-based safety navigation concept for senior citizens. In: 2nd International Conference on Applied and Theoretical Information Systems Research (2nd. ATISR2012), pp. 1–9 (2012)
42. Rae, I., Venolia, G., Tang, J.C., Molnar, D.: A framework for understanding and designing telepresence. In: Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing - CSCW 2015, Vancouver, BC, Canada, pp. 1552–1566. ACM Press (2015). http://dl.acm.org/citation.cfm?doid=2675133.2675141. https://doi.org/10.1145/2675133.2675141
43. Scheggi, S., Talarico, A., Prattichizzo, D.: A remote guidance system for blind and visually impaired people via vibrotactile haptic feedback. In: 22nd Mediterranean Conference on Control and Automation, pp. 20–23. IEEE (2014)
44. Sjöblom, M., Hamari, J.: Why do people watch others play video games? An empirical study on the motivations of twitch users. Comput. Hum. Behav. **75**, 985–996 (2017)
45. Slater, M., Linakis, V., Usoh, M., Kooper, R.: Immersion, presence and performance in virtual environments: an experiment with tri-dimensional chess. In: Proceedings of the ACM Symposium on Virtual Reality Software and Technology, VRST 1996, pp. 163–172. Association for Computing Machinery, New York (1996). https://doi.org/10.1145/3304181.3304216
46. Steltenpohl, H., Bouwer, A.: Vibrobelt: tactile navigation support for cyclists. In: Proceedings of the 2013 International Conference on Intelligent User Interfaces, pp. 417–426. ACM (2013)
47. Stewart, D.R., Littau, J.: Up, periscope: mobile streaming video technologies, privacy in public, and the right to record. Journal. Mass Commun. Q. **93**(2), 312–331 (2016)
48. Tang, A., Fakourfar, O., Neustaedter, C., Bateman, S.: Collaboration in 360 videochat: challenges and opportunities. Technical report, University of Calgary (2017)

49. Tsukada, K., Yasumura, M.: ActiveBelt: belt-type wearable tactile display for directional navigation. In: Davies, N., Mynatt, E.D., Siio, I. (eds.) UbiComp 2004. LNCS, vol. 3205, pp. 384–399. Springer, Heidelberg (2004). https://doi.org/10.1007/978-3-540-30119-6_23
50. Van Erp, J.B., Van Veen, H.A., Jansen, C., Dobbins, T.: Waypoint navigation with a vibrotactile waist belt. ACM Trans. Appl. Percept. (TAP) **2**(2), 106–117 (2005)
51. Wang, H.C., Katzschmann, R.K., Teng, S., Araki, B., Giarré, L., Rus, D.: Enabling independent navigation for visually impaired people through a wearable vision-based feedback system. In: 2017 IEEE International Conference on Robotics and Automation (ICRA), pp. 6533–6540. IEEE (2017)
52. Witmer, B.G., Singer, M.J.: Measuring presence in virtual environments: a presence questionnaire. Presence **7**(3), 225–240 (1998)