"Like I was There:" A User Evaluation of an Interpersonal Telepresence System Developed through Value Sensitive Design

KEVIN PFEIL, University of North Florida, USA
KARLA BADILLO-URQUIOLA, University of Notre Dame, USA
JOSEPH J. LAVIOLA JR., University of Central Florida, USA
PAMELA J. WISNIEWSKI, Vanderbilt University, USA

We developed and deployed an interpersonal telepresence prototype aimed at providing a positive one-to-one interaction between a Streamer and a Viewer. Our prototype uses four distributed, wearable cameras hidden from the public eye. It was designed to reduce the risk of Streamer self-consciousness while providing the Viewer with a greater sense of autonomy. We deployed our prototype with sixteen participants in dyads, who worked together to complete a scavenger hunt, and compared it the baseline of Skype. We found how our prototype better supported Streamer social well-being and physical comfort, and it also better supported Viewer autonomy. However, almost all participants desired a change to the design of the prototype, hinting that we need to provide better customization for future iterations of interpersonal telepresence devices.

CCS Concepts: • **Human-centered computing** → *HCI design and evaluation methods*; **User centered design**.

Additional Key Words and Phrases: Telepresence, User Evaluation, Qualitative Methods

ACM Reference Format:
Kevin Pfeil, Karla Badillo-Urquiola, Joseph J. LaViola Jr., and Pamela J. Wisniewski. 2024. "Like I was There:" A User Evaluation of an Interpersonal Telepresence System Developed through Value Sensitive Design. *Proc. ACM Hum.-Comput. Interact.* 8, CSCW2, Article 476 (November 2024), 18 pages. https://doi.org/10.1145/3687015

1 Introduction

The COVID-19 pandemic is now behind us, but it had many impacts on the global community — and we are now able to reflect upon the impact it had on mental health. For instance, social distancing, while necessary, prevented people from having social experiences and subsequently caused an increase in depression rates [4, 52]. Previous work suggests how richer social experiences may reduce depression [34], yet — as found in a study based in the United Kingdom — video chat is not as conducive to lowering rates of depression as face-to-face interaction [52]. This implies that our current technological solutions are insufficient, and the SIGCHI community must work towards fostering more meaningful social interactions using remote technologies.

One promising avenue that is increasingly popular among researchers is *Telepresence*, or the ability to perceive and/or interact with a remote environment as if actually there [36]. Though originally conceived for remote work, telepresence (in various forms) has been the subject of study in the

Authors' Contact Information: Kevin Pfeil, kevin.pfeil@unf.edu, University of North Florida, Jacksonville, FL, USA; Karla Badillo-Urquiola, kbadill3@nd.edu, University of Notre Dame, Notre Dame, IN, USA; Joseph J. LaViola Jr., jjl@cs.ucf.edu, University of Central Florida, Orlando, FL, USA; Pamela J. Wisniewski, pamela.wisniewski@vanderbilt.edu, Vanderbilt University, Nashville, TN, USA.

ACM 2573-0142/2024/11-ART476
https://doi.org/10.1145/3687015

context of social interaction. For instance, researchers have implemented novel telepresence devices in the home to support rich, remote interactions [16, 28, 59], and others have deployed telepresence for mobile activities [29, 30, 37, 50, 63]. This context — mobile, shared experiences between two individuals using live-streaming technologies — is referred to as *interpersonal telepresence*. This differs from typical live-streaming where one user streams to a large audience. This paradigm is valuable because it can be used to enhance relationships and provide a means to connect people who are in need; bedridden individuals (e.g. those in an extended hospital stay) can be provided with means beyond video chat to share rich experiences with their loved ones, which may improve surgical outcomes [31, 52]. Previous work in this space notes how the Streamer — the individual providing a remote experience for another — sometimes feels self-conscious using telepresence hardware in public (including simple devices such as smartphones), and the Viewer — the individual receiving a remote experience from the Streamer — may sometimes feel bored, disinterested, or disengaged, because the camera's view did not capture interesting stimuli, or because the camerawork of the Streamer was not conducive to a positive experience [12, 30, 47, 48, 50]. These problems have been identified in multiple phases of research, including the design (e.g. [47]) and evaluation (e.g. [12, 30]) phases, which hints that these constructs must be tackled directly. In a recent study that used Value-Sensitive Design (VSD) to identify the values that should be supported by future interpersonal telepresence prototypes, it was found how both major stakeholders, assuming they know each other, desired to have an interaction that accounted for both sides. In short, an ideal prototype would account for the Streamer's social well-being and physical comfort, and it would simultaneously allow the Viewer to have autonomy in their experience [47]. However, VSD methodology does not yield hardware; that step is left to future designers. In our paper, we undertake this role and create a prototype that we suspect directly embeds these values. Our prototype uses four wearable cameras embedded in clothing and accessories, with the intention that the Streamer will not stand out in public, and the Viewer could choose their own camera angle. Still, we must still evaluate our device to ensure it meets the needs of the end-users. Thus, with this work, we ask the following research questions:

RQ1: *How do interpersonal telepresence Streamers perceive a prototype aimed at supporting their social well-being with hidden cameras?*
RQ2: *How do interpersonal telepresence Viewers perceive a prototype aimed at supporting their autonomy through multiple camera views attached to the Streamer?*
RQ3: *How can we further improve interpersonal telepresence user experience?*

To answer these questions, we deployed our custom prototype that was intended to make Streamers less self-conscious while offering Viewers a sense of autonomy. Our developed prototype employs four wearable cameras that allows the Viewer to see in front of, to the left of, to the right of, and behind the Streamer. These cameras were sewn into the fabric of — and embedded into — a hat and a simple backpack. In this manner, the cameras were hidden as much as possible to the naked eye. To complement the physical hardware, we developed a custom WebRTC application that allowed the Viewer to switch between cameras with simple button presses. Our prototype was similar to a 360° camera, though it was hidden as much as possible from bystanders and passersby. We recruited sixteen participants in eight dyads comprised of friends or family, with whom we deployed our custom prototype as well as Skype, and we allowed them to use the systems in any manner they desired in pursuit of completing a scavenger hunt. We found how our Streamer participants generally preferred our prototype because it was comfortable, easy to use, and and allowed them to feel confident. Likewise, our Viewer participants generally preferred our prototype because there was more opportunity for exploration. However, both sides of the interaction offered suggestions to make the prototype even stronger. These recommendations were almost exclusively in regards to customization, implying the future of interpersonal telepresence is one that helps end-users shape the technology according to their own needs and desires. Our work makes the following contributions to the literature:

(1) We demonstrate how a successful interpersonal telepresence prototype can emerge from Value-Sensitive Design methodology.
(2) We highlight the importance of customization and echo plans for empowering end users in the context of interpersonal telepresence.

The following sections describe relevant literature which guided us in our development efforts, followed by our detailed design decisions and the final prototype which we created. After, we describe a user study in which we deployed our prototype, followed by the results of that study. We close with a discussion about the future of interpersonal telepresence.

2 Related Work

In this section, we review the relevant literature related to interpersonal telepresence. We begin with a brief history and state-of-the-art of telepresence, and we conclude with the gaps and opportunities that still need to be addressed with future work.

2.1 Evolution of Telepresence

In 1980, Minsky conceptualized telepresence as the remote control of robots using a first-person point of view to complete work [36]. The general idea was to use robotic platforms and other hardware that would stimulate users' senses in a natural way and to let them directly manipulate the remote environment, such that they would feel like they were "actually there". Today, robotic platforms with such affordances are not yet ubiquitous, but significant research has been conducted to bring forth consumer-grade mobile platforms including drones and wheeled robots [33]. Many barriers currently prevent robotic platforms from being accepted en mass, e.g. some people do not want to be close to them [15], they move too slow or cannot go everywhere desired [23, 43], or their use is restricted [1, 20], but we do expect robotic platforms to eventually be accepted in everyday life. In this work, however, we focus on the interpersonal telepresence paradigm, which does not involve a robotic platform. In this paradigm, a Streamer leverages mobile video chat and applies it in a way that fosters the sharing of experiences richer than simple chatting [48]. Now-ubiquitous video chat applications and platforms with mobile phone support — like Skype, FaceTime, Twitch.tv, Periscope, and YouTube — were introduced relatively recently, and an abundance of research has been conducted to understand how people can use these to share rich experiences either with large groups [21, 22, 38, 53, 55] or with a loved one [5, 9, 41, 42]. Such devices have been used or posited in research for various scenarios, including exploration and recreation [19, 24, 30, 54, 60–62], supporting relationships [3, 40], physical activities [49, 56], and shopping [9–11], among others, and the common theme is to let the Streamer, who is on-site, "bring" the Viewer with them via video chat or live-streaming technologies to share the experience.

A variety of projects have begotten various features that support this theme; for instance, Kratz et al. [32] used a shoulder-worn, actuated smartphone mount that lets the Viewer choose where to look by manipulating the physical hardware direction. Kasahara et al. [29] created the head-worn JackIn system, which uses multiple cameras stitched together to form a 360° view, such that a user can inspect the remote location using virtual reality. Cai and Tanaka [11] created a prototype in which the Streamer can see their partner's hand gestures by use of an Augmented Reality headset, and Yazaki et al. [61] developed a system that allows a Viewer to hop between multiple Streamers, each using a custom 360° camera rig. The features described by these works have helped us to understand how we can foster togetherness through interpersonal telepresence, but the presence of features is not immediately conducive to a positive experience. Our work demonstrates how a simple prototype devoid of advanced features can still meet the needs of end-users and provide a better experience than simple video chat, but in the next section, we underscore the constructs with which future efforts must grapple.

2.2 Gaps and Opportunities in Interpersonal Telepresence

In the Virtual Reality (VR) literature, one of the major constructs used for decades to measure the feeling of "being there" in a simulated environment is *Presence* [51, 57]. However, researchers have identified how presence may have different meanings to various individuals. For instance, one's feeling able to naturally manipulate or perceive the environment is not the same as one's feeling able to naturally communicate with other people. As such, the construct of presence can be divided into three main dimensions: *Spatial Presence*, *Social Presence*, and *Self-Presence* (the introduction in the work by Oh et al. [45] superbly distinguishes between these). Though these constructs stem from VR literature, they are applicable to the telepresence field, as the original concept posited manipulation and perception of the remote environment [36]. In a recent review of the literature, it was found how many projects apply various instruments to measure spatial presence, but few apply social presence questionnaires. As a whole, however, interpersonal telepresence researchers generally measure presence (in some form) as well as system usability and user satisfaction [48]. To enhance Viewer satisfaction and bring about greater senses of presence, more features are developed and added to the Streamer's equipment. For instance, the use of a 360° camera allows the Viewer to choose their own viewing angle at the cost of a misaligned experience [54], but Augmented Reality technology can reintroduce non-verbal gestures such as pointing [9], or mutually display the directions in which both Streamer and Viewer are looking [29, 39]. By adding these features which are often obtrusive or highly noticeable, Streamers may feel self-conscious. Studies have shown how some people simply do not feel comfortable using even the simplest devices, like smartphones [12, 30, 42, 48], hinting to a conflict between stakeholders. For instance, Rae et al. [50] and Kim et al. [30] used a body-worn solution in which a smartphone was mounted to a shoulder strap. Jones et al. [27] used a hand-held cameraphone, and Neustaedter et al. [42] used a variety of commodity configurations. Each of these projects were conducted in public settings, and participants reported feeling socially awkward or self-conscious because they stood out and felt that people were staring at them. In the literature, many recent papers have described advanced prototypes to enhance the viewing experience [48], and though self-consciousness has not generally been directly studied in these works, the devices are typically even more obtrusive and visible than smartphones. In a recent study that applied VSD methods, major values of *Streamer Social Acceptance* and *Viewer Autonomy* were explicitly desired by potential stakeholders, and due to the context of knowing the person on the other end of the interaction, both stakeholders expressed a desire to support the other [47]. Though that work was able to identify these values, it is a different challenge to identify hardware which can support them. In this work, we rise to this challenge by demonstrating how, by taking a value-sensitive approach, we were able to meet user needs, and strike a balance between both major stakeholders. In the next section, we describe the guidelines which emerged from previous VSD methodology. We also describe how we integrated values directly into our prototype's design, distinguishing it from existing solutions.

3 The Prototype Developed with Value-Sensitive Design Methods

Value-Sensitive Design is a theoretically grounded approach through which we can systematically embed values within the design process, using conceptual, empirical, and technical investigations [17]. In this section, we describe the design and development of an unobtrusive interpersonal telepresence prototype using VSD. Our approach was novel in that it addressed known gaps in the literature by integrating values directly in our design, which sets it apart from existing solutions in this space, which generally present features without consulting users in the design process. Our prototype's design was based on findings from a previous value elicitation study [47]; we summarize and paraphrase the guidelines of that work below:

- **Supporting the Streamer's Social Needs**: an ideal design would prevent the Streamer from feeling self-conscious while using the technology by being less bulky and noticeable by other people.
- **Reducing the Streamer's Burden (physical, financial, or otherwise)**: an ideal design would be lightweight, hands-free, unobtrusive, and straight-forward; there would be little effort to setup the interaction; and the implementation would be low-cost.
- **Granting the Viewer Freedom to Explore**: an ideal design would provide affordances to let them choose their own viewpoints, which would feel natural.

We aimed to incorporate as many of these values as possible into our prototype. We first focused on the Streamer's social needs; we wanted our prototype to be as inconspicuous as possible, so we elected to use hidden cameras embedded into clothing or accessories, reducing the overall visual footprint. To preserve privacy and reduce auditory noticeability, we planned to integrate earbuds instead of speakerphone. Next, to facilitate a physically comfortable experience, we confirmed the use of wearable devices, to free the Streamer's hands. To embed the low-cost value, we targeted the use of inexpensive, commercial off-the-shelf devices. Lastly, we aimed to ensure our prototype would let the Viewer choose any angle they desired. We wanted to apply a 360° camera, but this conflicted with the previous values. To maintain a low visual profile while still affording the Viewer with the capability to explore the environment on their own, we elected to use an array of cameras which would together provide a near-omnidirectional view. This decision was additionally intended to prevent occlusion by the wearer's body, which is known to detract from the user experience [46].

Our final interpersonal telepresence prototype thus involved the creation of custom Streamer hardware, and, since at the time of work there were no low-cost streaming platforms that would offer a solution for streaming multiple cameras simultaneously, custom software. The remainder of this section details the hardware specifications for our prototype as well as the software that supported the telepresence interaction.

3.1 The Streamer's Hidden Camera Rig

This section details the hardware chosen for our final Streamer setup. As previously mentioned, our intent was to create a wearable prototype that would not draw unwanted attention to the user while simultaneously providing multiple camera angles. We drew inspiration from previous work to guide our design. Kasahara et al. [29] created the JackIn system, which is a head-worn 360° device. While the capability provides exactly what we desired, there is currently no consumer-grade iteration of this device, and it seems to be too noticeable for our purposes [48]; to replicate functionality but reduce visibility and overall cost, we elected to embed four Logitech C270 Webcams into wearables. These webcams each had USB cables, were plug-and-play, and had a 60° field of view, for the low cost of $40 USD per device. We used an Xcellon 4-port Slim USB-C hub so they could all be used by our laptop, which was an Asus Strix G15. To stream the cameras over cellular network, we used a OnePlus Nord N200 as a hotspot.

In our first iteration, in which we embedded the cameras into the material of a sports cap, the electronics would heat up and also press against the user's head, causing discomfort. For our second iteration, we swapped the sports cap for a unisex boonie hat and sewed a ribbon around the outside. This way, the electronics would not make contact with the user's skin, and the ribbon was able to shroud the cameras. As the cameras we elected to use had USB cables that had to run from the

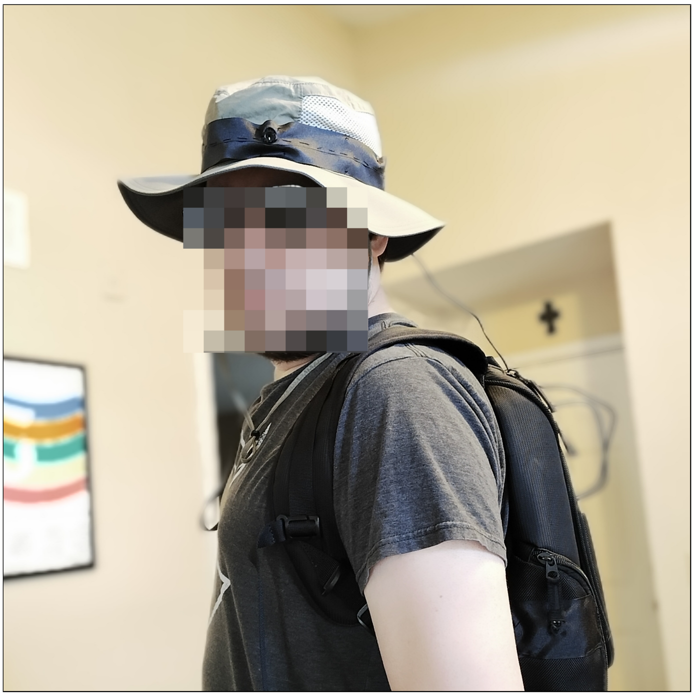

Fig. 1. The final prototype used in this work. We embedded a forward-facing camera into a boonie hat, and we embedded back, left, and right cameras into a backpack.

hat to the laptop, mounting all 4 cameras on the hat increased noticeability. Further, mounting cameras to the head meant all viewpoints would shift when the user rotated their head. To reduce the visibility of the cables and to stabilize the view, we moved three of the cameras to a backpack. These cameras were again embedded in the fabric, either on the left, side, or back of the backpack, and the cables were not visible. The final iteration of the Streamer's setup included a forward-facing camera embedded in the boonie hat as well as the other 3 cameras embedded in the backpack; see Figure 1 for illustration. While this setup does not create a complete 360° view, it echoes work of the past that used eye-mounted cameras ([2, 5, 49]) as well as those that used multiple cameras ([9, 30, 63]). In the next section we detail the custom software that allows the Viewer to freely explore the remote environment during an interpersonal telepresence experience.

3.2 The Viewer's Application

Having created the Streamer's wearable setup, we needed Viewer software that would complement it. Our custom setup utilized multiple webcams, and we anticipated how it might not be feasible to process and stream concurrent camera feeds across a cellular network from one source. Simultaneously, it possible to receive multiple feeds on platforms like Skype, but the Viewer would need a way to choose the feed they wanted to see, which currently is not an option. Thus, we needed to create our own platform such that the user can choose the feed they wanted to view. To achieve this, we developed a custom WebRTC[1] application that displayed one camera's feed on a large window. Next to the window, four buttons allowed the user to choose the camera feed

---

[1] http://webrtc.org

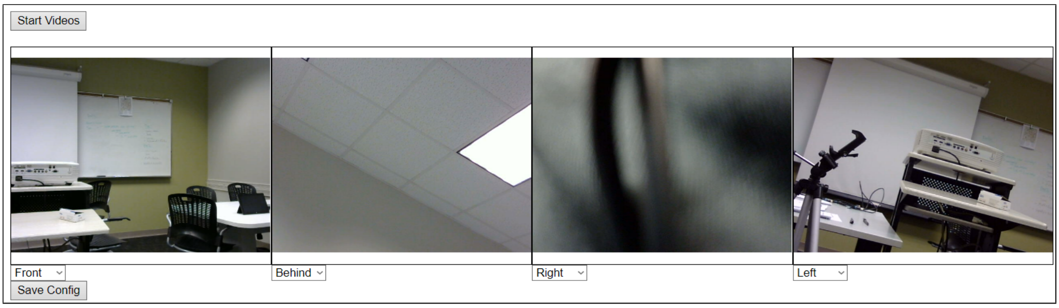

Fig. 2. A screen capture of the Streamer's client application. A user would simply need to determine which camera pointed in which direction using drop-down menus.

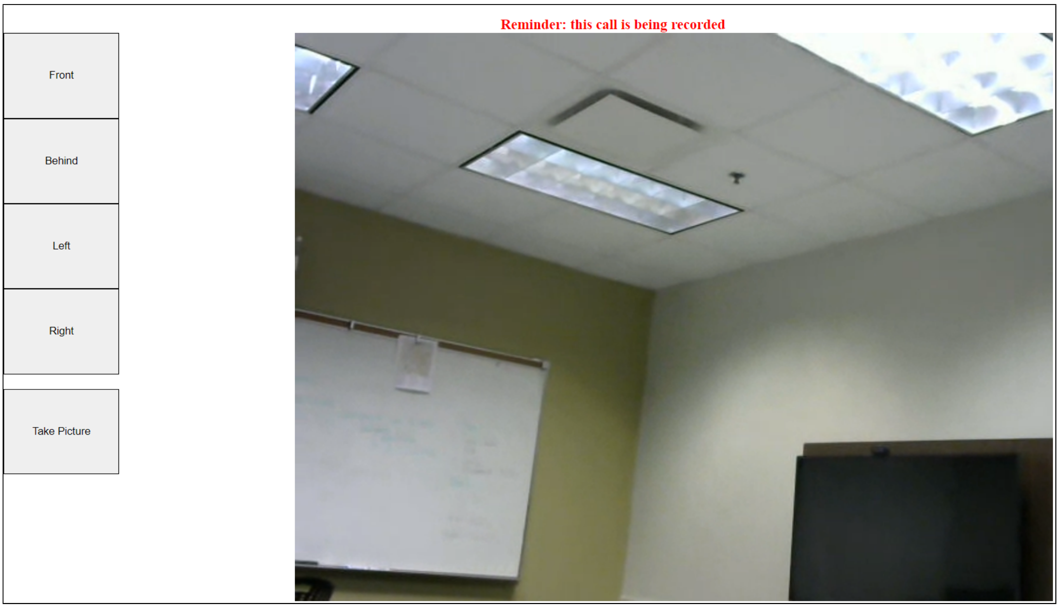

Fig. 3. A screen capture of the Viewer's client application. Buttons would allow a Viewer to choose the camera they want to view, as well as take a picture of the current camera view.

(front, behind, left, or right) they wanted to view. A fifth button was included to take a snapshot of the current view. We noticed that the WebRTC application did not detect the cameras in the same order each time (meaning the buttons did not always correspond to the correct camera). In order to ensure the buttons would bring up the correct video feed, we developed a simple module in the same WebRTC application that allowed the Streamer to pair each camera with a button. Therefore, when the user clicked one of the four directional buttons, the large window showing the feed would swap to the camera the user desired. See Figure 2 and Figure 3 for illustration. The web application was hosted on our Github Pages account.

With the front-end software created, we needed a way for the applications to communicate; i.e., we needed to send signals to the Streamer application whenever one of the buttons were pressed. To achieve this, we needed to use a STUN server, a TURN server, and a Signalling server, as denoted in the WebRTC documentation. Since our solution was in essence a peer-to-peer solution, we needed a way to ensure the data packets would successfully transmit between both end user devices. We utilized the publicly-available STUN server hosted by Google, and requisitioned an Amazon Web Services Elastic Compute Cloud (AWS EC2) server such that we could deploy a Coturn server[2]. We followed a set of existing instructions that guided us through the process (see reference [35]). Finally, we deployed a Node.js Signalling server on Heroku to send messages between the two end users. For instance, when a directional button was pressed on the Viewer's application, a message was sent to the signalling server, which relayed the message to the Streamer's application; this message then triggered the appropriate response, which was to shut off the current camera and turn on the one corresponding to the button press.

4 Methods

In this section, we describe a user study in which we deployed our prototype. We recruited dyads who knew each other to participate in our study by using our custom prototype and Skype with a smartphone. We chose Skype because it has one-to-one, bidirectional video chat features commonly studied in previous work (e.g. [3, 25, 30, 42]), and a handheld smartphone with video chat is commodity technology that has become a mainstay in daily life.

4.1 Study Design

We used a 2x1 within-subjects design for our study, and the only independent variable was *Device*; the two conditions were *Skype* and *Prototype*. We did not have any dependent variables, as our study is qualitative in nature. We randomly assigned condition order and applied counter-balancing to reduce bias. Our study was approved by our institution's IRB.

4.2 Subjects

As our work is in the interpersonal telepresence context, we recruited 10 user dyads to participate in our study. One pair did not appear for their scheduled time, and another was canceled due to a severe thunderstorm. Our final participant pool consisted of 8 user dyads (N = 16). Within each dyad, the participants were required to be friends or family, i.e. they needed to know each other very well. We asked our participants their age, gender, and relationship with their partner, and we also administered the Mini-IPIP [14], an inventory that measures the Big Five personality traits (Openness, Conscientiousness, Extraversion, Agreeableness, and Neuroticism [18, 26]) with a smaller set of questions. See Table 1 for participant demographics.

4.3 Apparatus

For the custom camera rig condition, the Streamer was given the developed prototype (previously described). To support audio communication, we used Jabra Elite 3 earbuds that connected to the laptop via BlueTooth; they included an embedded microphone, and we enabled the pass-through audio mode such that the user could hear their surroundings. The Viewer participants received the stream on a Microsoft Surface tablet that had an Intel i5 processor clocked at 1.7GHz, 4GB of RAM, and the Windows 8 operating system; this device had a built-in microphone such that they could verbally communicate with the Streamer. For the Skype condition, we asked the streaming participant to use their own phone; but in the event that this was not possible, we let them use our OnePlus Nord N200. The viewer participant viewed the stream on the same Surface tablet for both conditions. Both participants simply used the basic features of Skype, which supports audio communication and video streaming from a laptop's webcam or smartphone's built-in camera.

---

[2] https://github.com/coturn/coturn

Table 1. Participant Demographics. We recruited eight unique participant dyads, who chose their own roles. We provide age, gender, and Big 5 personality trait measurements: Openness, Conscientiousness, Extroversion, Agreeableness, and Neuroticism.

ID	Relationship	Role	Age	Gender	O	C	E	A	N
1	Married	Streamer	27	F	3.8	2.0	3.0	3.0	3.3
		Viewer	30	M	3.0	4.0	3.0	3.8	3.5
2	Friends	Streamer	22	M	3.0	3.3	2.5	4.3	3.5
		Viewer	30	M	5.0	3.8	3.5	4.3	2.0
3	Friends	Streamer	25	M	4.3	4.5	3.5	4.0	3.3
		Viewer	22	M	3.3	3.8	2.5	4.0	2.3
4	Partners	Streamer	23	M	4.3	4.0	3.3	4.0	2.0
		Viewer	22	F	4.3	2.8	4.3	3.5	3.3
5	Friends	Streamer	26	M	3.5	3.0	1.8	3.5	3.0
		Viewer	20	M	4.0	3.0	2.8	4.3	2.3
6	Best Friends	Streamer	19	F	5.0	2.8	2.5	3.8	4.0
		Viewer	20	F	4.8	3.8	3.0	4.0	1.5
7	Classmates	Streamer	24	M	3.5	4.0	4.0	4.5	2.5
		Viewer	23	M	3.8	2.3	2.3	3.5	2.5
8	Friends	Streamer	19	M	3.8	4.0	2.8	3.3	1.8
		Viewer	19	M	3.5	3.3	3.3	4.0	2.8

4.4 Procedure

The study occurred at the University of Central Florida's main campus in Orlando, Florida, where there are many buildings with art installations scattered about the premises. UCF's campus is circular in shape and about 5.7 square kilometers in size, with a major road enveloping the majority of the buildings. Our participant dyads met us inside one building located near one edge of the premises, and after completing the informed consent process, we issued a demographics questionnaire. Following, we explained that one participant would need to act as a Streamer for both study conditions, while the other would need to act as the Viewer. We let the participants decide how they wanted to assume their respective roles, but we explained how they cannot switch roles during the study. After they determined the roles, we collected photo IDs from both participants, which we held until the study was complete. We then applied the first condition in a randomized, counter-balanced order. For the Skype condition, we had the Streamer participant call the researcher's Surface tablet, and we received the call using the researcher's Skype account. For the Prototype condition, we showed the participants how the devices operated, and then had the Streamer don the hat, backpack, and wireless ear buds.

We gave the Streamer a map of the campus, and sent them to complete the task, described below. When the Streamer left the room, we gave the Viewer a map which was marked with artwork locations. In addition, we supplied the Viewer with a simple website that displayed the artwork that was needed for the current task iteration. In this way, the Viewer knew which artwork was needed, and where exactly they resided, while the Streamer only knew the general layout of the premises. We recorded the interaction with a smartphone camera, pointing it at the Viewer and their screen. After the task was completed, or after 60 minutes elapsed, the Streamer was recalled back to the starting location. We allowed the Streamer to have a break, and when they were ready, we repeated the procedure with the second condition. When the participants completed the experiment, we conducted a semi-structured interview to collect detailed feedback about the conditions. The study

Table 2. Semi-structured interview questions used in our study.

(Streamer) Which was easier to use and why?
(Viewer) How could we have improved your experience?
Why did you choose the roles that you chose?
Between the two systems, which did you like better and why?
Which made you felt more socially connected and why?
Which made you feel more self-conscious and why?
Which was more useful and engaging with the environment and why?
If you could redesign the Streamer capabilities, how would you do that?
Were there any other considerations we should have taken?
What type of situation would you use this system and why?
When would it not be useful?
For what context do you think Skype would work better?
For what context do you think the system would work better?

lasted approximately three hours, and at the end of the session, each participant was compensated $30 USD in cash.

4.5 Scavenger Hunt Task

For our study, the participants needed to work together to complete an artwork scavenger hunt. Around the premises, there were many art installations and paintings, both outside and indoors. We selected a subset of these artworks to include in our scavenger hunt task, favoring the outdoor art to prevent network connectivity issues; in total, we selected 14 pieces of art scattered around campus, but divided these into two subsets. We did not polarize the artwork, instead choosing to have the locations roughly overlap; the presented art subsets were counterbalanced along with condition. To complete the scavenger hunt, the participants needed to work together such that the Viewer would guide the Streamer to each piece of art so that the Viewer could take a picture of it. The picture needed to closely match the example picture provided to the Viewer. Our task is similar to that of Tang et al., who used a 360° camera mounted on a pole worn by the Streamer, who needed to take pictures of art from certain angles described by the Viewer [54]. One difference is that our Viewers were tasked with taking pictures of the art instead of the Streamer.

4.6 Semi-Structured Interview Questions

After both conditions were completed, we finished the session by administering a semi-structured interview with both participants simultaneously. We used the questions in Table 2 as a guide for our semi-structured interview, asking follow-up questions to delve deeper into user feedback to help uncover meaning and more details. Interviews lasted on average between 20-30 minutes. All interviews were audio recorded and transcribed by Otter.ai for analysis.

4.7 Data Analysis Approach

We conducted a thematic analysis [13] to understand the user experiences of Streamers (perceived confidence) and Viewers (perceived autonomy) with interpersonal telepresence. The first author started by reading through the transcriptions and using open coding to generate the first iteration of codes. The remaining authors were consulted to form consensus on the codes and subsequently throughout the iterative process. Axial coding was used to merge and organize related codes into themes. In our findings, we report the major themes that emerged from our dataset. When providing illustrative quotations, we combine the session ID with the participant role; for instance, S1 corresponds to the Streamer from session 1, and V8 corresponds to the Viewer from session 8.

5 Results

In this section, we present the results of our qualitative analysis. In general, participants on both sides of the interaction preferred our prototype to Skype video chat, offering how it was conducive to a more comfortable Streaming experience and how there were more Viewing opportunities. However, there are many opportunities to improve our design even further. We first provide details about how the participants communicated with each other, how they tackled the scavenger hunt task, and how long they took to complete the task.

5.1 Descriptive Findings

We recall how we instructed participants to complete the task in any manner they saw fit. The Streamers did just this — 4 of them streamed while on foot, with 1 of these choosing to run; 3 streamed while riding a bike to travel, and the final 1 used a longboard to travel. All Streamers used the same method to navigate for both conditions. Only one dyad was unable to complete the scavenger hunt task within the allotted 60 minutes, and this was during the Prototype condition. The participant pair had difficulty communicating, and the Streamer accidentally misaligned the camera, meaning their "forward" camera faced the wrong direction. At the 60-minute mark, that participant was recalled, but it took them an additional 13 minutes to return to the starting point. The average time taken to complete the task when Streaming with the prototype was 46.63 minutes (min: 28; max: 73), and the average time for Skype was 44.38 minutes (min: 36; max: 58). Regarding communication between participants, participants uttered, on average, 2175 words during the Prototype condition (min: 722; max: 4225), and they uttered, on average, 1813 words in the Skype condition (min: 992; max: 3044).

5.2 RQ1: Streamers Feel More Comfortable with our Prototype than when Using a Smartphone with Video Chat

We next report our findings surrounding the Streamer feedback. When choosing the preferred device, six of the eight Streamers selected our wearable prototype over Skype. There were multiple reasons behind this selection; the first reason that was found the most in our dataset was because their **hands were free** (5/8). The participants who rode their bike during the task were particularly satisfied with our prototype, because they needed to focus on steering; having a hands-free system allowed them to skip many steps when communicating with their partner:

> *"I definitely prefer [the prototype]. I was riding a bicycle, so [with Skype] I couldn't use my phone in my hand all the time. If I got to ask him, 'Am I at the right place?', I got to stand up from the bicycle, take the phone from my pocket, and show him. It seems like with Skype, I felt like we lost a degree of freedom in our body."* — S2

As intended, the prototype was conducive to a socially acceptable streaming experience that allowed the participants to feel **less self-conscious** (4/8). This contrasts to when the participants were using Skype; although video chat with a mobile phone is ubiquitous, it does draw attention to the user — and our participants felt "weird" recording the environment, including other people, with a gesture that brought unwanted attention:

> *"I definitely felt more self-conscious holding the phone, because there were people at these art installations working on them, and I'm literally just recording them so bad. Like, I promise I'm not stalking you guys, or this isn't going anywhere weird!"* — S6

Lastly, our Streamer participants — especially those who were on foot — preferred the prototype to mobile video chat because it was **easier to use**; this was discussed by 3 of our participants. For instance, one participant noted how, since the prototype had multiple camera angles, they didn't need to perform any camera-work; instead, the Viewer had the ability to accomplish any exploration by themselves, thus sharing the workload:

> *"[With the prototype] you can actually get a whole scope of everything that's going on instead of me having to do this [acts like they were manipulating a cameraphone]. And, maybe I miss something, so you don't actually get the full experience."* —S4

Thus, our Streamer participants generally preferred our prototype which made use of features to support their social well-being and physical comfort. In the next section, we report findings pertaining to the Viewer.

5.3 RQ2: Viewers Enjoy Autonomy when Exploring Remote Environments

We now report our findings surrounding Viewer feedback. When choosing the preferred device, six of the eight Viewers selected our wearable prototype to Skype. The main reason for selecting our prototype was because there was clearly **greater opportunity to explore** (6/8). We often observed our Viewer participants giving directions to the Streamers, and having more than one camera angle allowed our participants to get a better understanding of the remote environment. In downtime, i.e. while the Streamer was navigating between destinations, the Viewers typically engaged with the application, switching between the different cameras:

> *"I liked watching [with the prototype] because in between, like passing periods... I could have more of an idea of where he is. With the phone, you will you obviously want to walk around like this [acts like they were manipulating a cameraphone]. Here, I enjoyed watching the [prototype] one more."* — V8

For the dyads in which the Streamer rode a bike, this sentiment was even greater. When using Skype, those individuals had a difficult time multitasking; they often resorted to stowing their phone in their pocket as they navigated, only taking out the device when stopping for directions or completing the scavenger hunt task. Though the partners could still talk to each other, the downtime was often filled with silence. This completely hindered the ability for the Viewer to explore and detracted from the overall experience; thus, our hands-free prototype succeeded in providing the Viewer with more opportunities to explore:

> *"I feel like the [prototype] one is better, because with the one with Skype, he tends to put his phone away into one of his pockets, so I couldn't see anything at the time. So yeah, [the prototype] is more interactive."* — V5

In general, both Streamer and Viewer participants preferred the hands-free, multi-camera prototype compared to Skype, and the reasons were rather straightforward. However, the prototype was not without its problems. We received constructive feedback from both stakeholders; in the next section, we detail opportunity to improve upon our prototype.

5.4 RQ3: Both Streamers and Viewers Desire Customizability

In addition to the positive feedback, we received constructive criticism that is applicable to our prototype's design. Firstly, almost every Streamer participant (7/8) voiced how, if they would use a similar system again, they would **prefer a different style**. Mainly, the unisex boonie hat was met with disapproval. Some participants felt how the choice of hat affected their social well-being, because it not the type of hat they would normally wear. As participant S3 summarizes, *"appearance matters."* From the Viewer's perspective, the choice of hat also had an effect on usability. For instance, the individuals riding bicycles ran into a problem where the brim of the hat would flip up due to the wind, and this completely obstructed the view. Further, since the cameras were wearable, they were prone to capturing the environment from somewhat strange angles based on the orientation of their wearer's body. Our Streamer participants had various walking gaits, and some of them leaned forward while they walked. Others naturally tilted their heads, meaning the front camera was sometimes misaligned with the direction they were traveling. Additionally, one of the female participants had long hair, which got in the way of the Backwards camera. Although we encountered these problems, our participants were enthusiastic about using wearable devices, offering suggestions to improve their experience. These suggestions included using cameras with wider fields of view, which will naturally enhance the ability to explore; and, others offered different wearable solutions. For instance, some suggested mounting a camera on the Streamer's shirt or jacket, to decouple the view from the head. While this may prevent the Viewer from knowing exactly where the Streamer looks, it is a solution that would centralize the cameras and perhaps stabilize the views.

6 Discussion

In this section, we provide a discussion of our findings. We also help to map the future of interpersonal telepresence research.

6.1 Value-Sensitive Design Methods Helped Yield a Successful Prototype

In a 2021 review of the interpersonal telepresence literature, it was found how most of the work in this space since 2010 was technology-focused, i.e. the development of features has been paramount [48]. Many of the articles in that review presented an evaluation of new features in a controlled environment, and the respective participants gave valuable feedback to the particular devices. One of our critiques for such articles is that the insights are limited to the artifacts, or the particular classification of those artifacts. In contrast, our work follows VSD methodology, which does not focus on features; rather, it focuses on the values that can be embedded within features [6, 17]. As our paper used the results of a previous value elicitation study that found how Streamer social well-being, physical comfort, and Viewer autonomy are values for which we must account [47], the results in our paper do not regard features; rather, they regard the *techniques* we used to encompass the values. The techniques we used for our prototype — embedding multiple cameras into clothing and accessories — was satisfactory, highlighting how VSD methods are appropriate in this domain. One strength of this approach is that, as hardware for VR and 3D user interfaces continues to evolve, the identified values will still be need to be considered; techniques such as the ones we present here may outlast the features. Thus, we do not provide design recommendations through our work; instead, we provide a potential solution that is comprised of various techniques. We find how hiding the camera into wearables is *one of many* solutions, and through this solution, we can reduce Streamer self-consciousness and support Viewer engagement. However, as we discuss in the next section, this solution may not work for all individuals — we must tackle individual differences, and we can do so through supporting customization.

6.2 One Size Does Not Fit All: Towards Customizable Interpersonal Telepresence

Interpersonal telepresence research has resulted in many devices to support this paradigm [48]. As above, a critique for the literature is how many of these works use quantitative methods to show how, statistically, one feature is superior to another, through one or many metrics. It is important to note how with these methods, we evaluate our features with a group of individuals — yet it is not the group for whom we are designing: rather, we are aiming to support individuals. We can learn from our work how, even though we were able to support the Streamer's social well-being and physical comfort with our technique, there is still much to be desired. In this case, many users want to have a device that is customizable, i.e. a way to modify the fashion. Since we did not build any technology and instead used commercial off-the-shelf products, it is reasonable to suggest that we can *teach* users of interpersonal telepresence to build or configure their own devices. This is similar to previous human-computer interaction work, in which researchers created toolkits and instructed end-users how to build their own electronics. For example, we are motivated by Buechley and Eisenberg [7], who created "e-textiles" with which novices were able to build their own wearable electronics [7, 8]. Our work, following VSD methods, suggests how we need to empower our users into creating their own interpersonal telepresence devices. We envision an interpersonal telepresence toolkit, with which novice users can learn the skills to embed their devices into clothing and other wearable accessories, such that they would feel *as comfortable as possible* with their design. There is a psychological phenomenon called the "IKEA effect", in which people value the products that they build [44]; it thus follows that Streamers may be more satisfied with a telepresence prototype if they have input in the creation step.

6.3 Limitations and Future Work

Our study is constrained by several limitations that should be addressed in future work. First, the task duration was relatively lengthy, and as it was a within-subjects design, there was certainly a physical toll exacted from our Streamers. We did not measure physical attributes of our users, but we do suspect that the weight of our prototype influenced responses; for instance, one participant ran around campus in a display of physical aptitude, while others walked and required a substantial break before the second leg of the trial. We also acknowledge that our results cannot generalize to every possible usage scenario. Our study was conducted outside, during sunny weather, in the heat; we cannot say for certain that our devices would be appropriate indoors, or in cold or rainy weather. We would expect that a prototype such as ours would be met with much hesitation during more formal settings, or in private spaces. As such, additional work should be conducted with alternative venues and use cases in mind.

Second, our findings may have been influenced by sampling bias and/or a limited sampling frame. While we measured Big Five Personality Traits [18, 26] in our pre-survey in anticipation of seeing individual differences based on personality, most of our participants fell within the central tendency of this scale, and we did not have enough statistical power to detect significant differences. Therefore, future work could conduct a similar study with a larger sample size and more quantitative measures to determine if personality type mediates or moderates user experience. We believe that it will, as we saw a trend where the more extroverted individual (from observation) tended to be the person who opted to be the streamer role. However, future work should confirm this qualitative trend through more quantitative and longitudinal means.

Finally, We also did not capture any data from the third telepresence stakeholder — the bystanders who are collocated in the environment with the Streamer. The technique used in our work involves hiding a camera from other people, which underscores the need for ethical discussion. While people in countries like the United States have a right to record any public place without gathering consent from people within that space, we do not want to conflate legality with ethics. With this paper, we are not arguing for or recommending the use of hidden camera technology in private spaces; and regarding public spaces, further discussion must be held to balance the Streamer's social needs with bystander's privacy [58]. Future work will help capture data directly from bystanders to understand how we can better accommodate them within the design process.

7 Conclusion

In this work, we describe an interpersonal telepresence prototype that successfully embedded values identified by previous work. By embedding cameras inside clothing and accessories, we were able to achieve a low-profile prototype that still offered more opportunity to explore a remote environment than the baseline of video chat. However, there is still much work to be done towards supporting Streamer individual differences. We find how our prototype is less obtrusive than a smartphone, but it is not usable in every situation. Our work is specific to shared, outdoor experiences in an informal setting, but individual differences directing design decisions need to be considered in future iterations of prototypes such as ours. To our knowledge, our work describes the first interpersonal telepresence prototype developed using VSD, and we expect future work to present additional efforts to support the stakeholders in this context.

References

[1] 2020. Certificated Remote Pilots including Commercial Operators. https://www.faa.gov/uas/commercial_operators/

[2] Sudhanshu S.D.P. Ayyagari, Kunal Gupta, Matt Tait, and Mark Billinghurst. 2015. CoSense: Creating Shared Emotional Experiences. In *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems - CHI EA '15*. ACM Press, Seoul, Republic of Korea, 2007–2012. https://doi.org/10.1145/2702613.2732839

[3] Uddipana Baishya and Carman Neustaedter. 2017. 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 - CSCW '17*. ACM Press, Portland, Oregon, USA, 84–97. https://doi.org/10.1145/2998181.2998200

[4] Alexandra Benisek. 2021. *COVID-19 and Depression.* Retrieved January 1, 2024 from https://www.webmd.com/covid/covid-19-depression

[5] Mark Billinghurst, Alaeddin Nassani, and Carolin Reichherzer. 2014. Social panoramas: using wearable computers to share experiences. In *SIGGRAPH Asia 2014 Mobile Graphics and Interactive Applications on - SA '14*. ACM Press, Shenzhen, China, 1–1. https://doi.org/10.1145/2669062.2669084

[6] Alan Borning and Michael Muller. 2012. Next Steps for Value Sensitive Design. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Austin, Texas, USA) *(CHI '12)*. Association for Computing Machinery, New York, NY, USA, 1125–1134. https://doi.org/10.1145/2207676.2208560

[7] Leah Buechley and Michael Eisenberg. 2008. The LilyPad Arduino: Toward Wearable Engineering for Everyone. *IEEE Pervasive Computing* 7, 2 (2008), 12–15. https://doi.org/10.1109/MPRV.2008.38

[8] Leah Buechley, Mike Eisenberg, Jaime Catchen, and Ali Crockett. 2008. The LilyPad Arduino: Using Computational Textiles to Investigate Engagement, Aesthetics, and Diversity in Computer Science Education. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Florence, Italy) *(CHI '08)*. Association for Computing Machinery, New York, NY, USA, 423–432. https://doi.org/10.1145/1357054.1357123

[9] Minghao Cai, Soh Masuko, and Jiro Tanaka. 2018. Gesture-based Mobile Communication System Providing Side-by-side Shopping Feeling. In *Proceedings of the 23rd International Conference on Intelligent User Interfaces Companion - IUI 18*. ACM Press, Tokyo, Japan, 1–2. https://doi.org/10.1145/3180308.3180310

[10] Minghao Cai and Jiro Tanaka. 2017. Trip Together: A Remote Pair Sightseeing System Supporting Gestural Communication. In *Proceedings of the 5th International Conference on Human Agent Interaction - HAI '17*. ACM Press, Bielefeld, Germany, 317–324. https://doi.org/10.1145/3125739.3125762

[11] Minghao Cai and Jiro Tanaka. 2019. Go together: providing nonverbal awareness cues to enhance co-located sensation in remote communication. *Human-centric Computing and Information Sciences* 9, 1 (Dec. 2019), 19. https://doi.org/10.1186/s13673-019-0180-y

[12] Anezka Chua, Azadeh Forghani, and Carman Neustaedter. 2017. Shared Bicycling Over Distance. In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems - CHI EA '17*. ACM Press, Denver, Colorado, USA, 455–455. https://doi.org/10.1145/3027063.3049776

[13] Victoria Clarke and Virginia Braun. 2021. *Thematic analysis: a practical guide.* SAGE Publications Ltd.

[14] M Brent Donnellan, Frederick L Oswald, Brendan M Baird, and Richard E Lucas. 2006. The mini-IPIP scales: tiny-yet-effective measures of the Big Five factors of personality. *Psychological assessment* 18, 2 (2006), 192.

[15] Brittany A. Duncan and Robin R. Murphy. 2017. Effects of Speed, Cyclicity, and Dimensionality on Distancing, Time, and Preference in Human-Aerial Vehicle Interactions. *ACM Transactions on Interactive Intelligent Systems* 7, 3 (Sept. 2017), 1–27. https://doi.org/10.1145/2983927

[16] Sean Follmer, Hayes Raffle, Janet Go, Rafael Ballagas, and Hiroshi Ishii. 2010. Video Play: Playful Interactions in Video Conferencing for Long-distance Families with Young Children. In *Proceedings of the 9th International Conference on*

Interaction Design and Children (Barcelona, Spain) *(IDC '10)*. ACM, New York, NY, USA, 49–58. https://doi.org/10.1145/1810543.1810550

[17] Batya Friedman, Peter H. Kahn, Alan Borning, and Alina Huldtgren. 2013. *Value Sensitive Design and Information Systems*. Springer Netherlands, Dordrecht, 55–95. https://doi.org/10.1007/978-94-007-7844-3_4

[18] Lewis R Goldberg. 1990. An alternative" description of personality": the big-five factor structure. *Journal of personality and social psychology* 59, 6 (1990), 1216. Publisher: American Psychological Association.

[19] Lilian de Greef, Meredith E. Morris, and Kori Inkpen. 2016. TeleTourist: Immersive Telepresence Tourism for Mobility-Restricted Participants. In *Proceedings of the 19th ACM Conference on Computer Supported Cooperative Work and Social Computing Companion - CSCW '16 Companion*. ACM Press, San Francisco, California, USA, 273–276. https://doi.org/10.1145/2818052.2869082

[20] Michael Gwilliam. 2020. Japanese Twitch streamer reveals the problem with IRL streaming in Tokyo. https://www.dexerto.com/entertainment/japanese-twitch-streamer-reveals-the-problem-with-irl-streaming-tokyo-1325147

[21] Oliver L. Haimson and John C. Tang. 2017. What Makes Live Events Engaging on Facebook Live, Periscope, and Snapchat. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. ACM, Denver Colorado USA, 48–60. https://doi.org/10.1145/3025453.3025642

[22] William A. Hamilton, John Tang, Gina Venolia, Kori Inkpen, Jakob Zillner, and Derek Huang. 2016. Rivulet: Exploring Participation in Live Events through Multi-Stream Experiences. In *Proceedings of the ACM International Conference on Interactive Experiences for TV and Online Video - TVX '16*. ACM Press, Chicago, Illinois, USA, 31–42. https://doi.org/10.1145/2932206.2932211

[23] Yasamin Heshmat, Brennan Jones, Xiaoxuan Xiong, Carman Neustaedter, Anthony Tang, Bernhard E. Riecke, and Lillian Yang. 2018. 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 - CHI '18*. ACM Press, Montreal QC, Canada, 1–13. https://doi.org/10.1145/3173574.3173933

[24] Kori Inkpen, Brett Taylor, Sasa Junuzovic, John Tang, and Gina Venolia. 2013. Experiences2Go: sharing kids' activities outside the home with remote family members. In *Proceedings of the 2013 conference on Computer supported cooperative work*. ACM, 1329–1340.

[25] Clarissa Ishak, Carman Neustaedter, Dan Hawkins, Jason Procyk, and Michael Massimi. 2016. Human Proxies for Remote University Classroom Attendance. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems - CHI '16*. ACM Press, Santa Clara, California, USA, 931–943. https://doi.org/10.1145/2858036.2858184

[26] Oliver P John, Sanjay Srivastava, et al. 1999. The Big-Five trait taxonomy: History, measurement, and theoretical perspectives. (1999).

[27] Brennan Jones, Anna Witcraft, Scott Bateman, Carman Neustaedter, and Anthony Tang. 2015. Mechanics of Camera Work in Mobile Video Collaboration. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems - CHI '15*. ACM Press, Seoul, Republic of Korea, 957–966. https://doi.org/10.1145/2702123.2702345

[28] Tejinder K. Judge, Carman Neustaedter, and Andrew F. Kurtz. 2010. The Family Window: The Design and Evaluation of a Domestic Media Space. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Atlanta, Georgia, USA) *(CHI '10)*. ACM, New York, NY, USA, 2361–2370. https://doi.org/10.1145/1753326.1753682

[29] Shunichi Kasahara, Shohei Nagai, and Jun Rekimoto. 2017. JackIn Head: Immersive Visual Telepresence System with Omnidirectional Wearable Camera. *IEEE Transactions on Visualization and Computer Graphics* 23, 3 (March 2017), 1222–1234. https://doi.org/10.1109/TVCG.2016.2642947

[30] Seungwon Kim, Sasa Junuzovic, and Kori Inkpen. 2014. 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 '14*. ACM Press, Sanibel Island, Florida, USA, 167–177. https://doi.org/10.1145/2660398.2660409

[31] Henning Krampe, Anke Barth-Zoubairi, Tatjana Schnell, Anna-Lena Salz, Léonie F Kerper, and Claudia D Spies. 2018. Social relationship factors, preoperative depression, and hospital length of stay in surgical patients. *International journal of behavioral medicine* 25 (2018), 658–668.

[32] Sven Kratz, Daniel Avrahami, Don Kimber, Jim Vaughan, Patrick Proppe, and Don Severns. 2015. 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 - MobileHCI '15*. ACM Press, Copenhagen, Denmark, 567–575. https://doi.org/10.1145/2786567.2787134

[33] Annica Kristoffersson, Silvia Coradeschi, and Amy Loutfi. 2013. A Review of Mobile Robotic Telepresence. *Advances in Human-Computer Interaction* 2013 (2013), 1–17. https://doi.org/10.1155/2013/902316

[34] G. Lewis, D.-Z. Kounali, K. S. Button, L. Duffy, N. J. Wiles, M. R. Munafò, C. J. Harmer, and G. Lewis. 2017. Variation in the recall of socially rewarding information and depressive symptom severity: a prospective cohort study. *Acta Psychiatrica Scandinavica* 135, 5 (2017), 489–498. https://doi.org/10.1111/acps.12729 arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1111/acps.12729

[35] Kostya Malsev. 2021. *Set up a TURN server on AWS in 15 minutes.* Retrieved March 8, 2022 from https://kostya-malsev.medium.com/set-up-a-turn-server-on-aws-in-15-minutes-25beb145bc77

[36] Marvin Minsky. 1980. Telepresence. Omni Magazine. *New York, Jun* (1980).

[37] Kana Misawa and Jun Rekimoto. 2015. Wearing another's personality: a human-surrogate system with a telepresence face. In *Proceedings of the 2015 ACM International Symposium on Wearable Computers - ISWC '15*. ACM Press, Osaka, Japan, 125–132. https://doi.org/10.1145/2802083.2808392

[38] Ahmed E. Mostafa, Kori Inkpen, John C. Tang, Gina Venolia, and William A. Hamilton. 2016. SocialStreamViewer: Guiding the Viewer Experience of Multiple Streams of an Event. In *Proceedings of the 19th International Conference on Supporting Group Work - GROUP '16*. ACM Press, Sanibel Island, Florida, USA, 287–291. https://doi.org/10.1145/2957276.2957286

[39] Shohei Nagai, Shunichi Kasahara, and Jun Rekimoto. 2015. LiveSphere: Sharing the Surrounding Visual Environment for Immersive Experience in Remote Collaboration. In *Proceedings of the Ninth International Conference on Tangible, Embedded, and Embodied Interaction - TEI '14*. ACM Press, Stanford, California, USA, 113–116. https://doi.org/10.1145/2677199.2680549

[40] Carman Neustaedter and Saul Greenberg. 2012. Intimacy in long-distance relationships over video chat. In *Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems - CHI '12*. ACM Press, Austin, Texas, USA, 753. https://doi.org/10.1145/2207676.2207785

[41] Carman Neustaedter, Carolyn Pang, Azadeh Forghani, Erick Oduor, Serena Hillman, Tejinder K. Judge, Michael Massimi, and Saul Greenberg. 2015. Sharing Domestic Life through Long-Term Video Connections. *ACM Transactions on Computer-Human Interaction* 22, 1 (Feb. 2015), 1–29. https://doi.org/10.1145/2696869

[42] Carman Neustaedter, Jason Procyk, Anezka Chua, Azadeh Forghani, and Carolyn Pang. 2020. Mobile Video Conferencing for Sharing Outdoor Leisure Activities Over Distance. *Human-Computer Interaction* 35, 2 (March 2020), 103–142. https://doi.org/10.1080/07370024.2017.1314186

[43] Carman Neustaedter, Gina Venolia, Jason Procyk, and Dan Hawkins. 2016. 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 - CSCW '16*. ACM Press, San Francisco, California, USA, 417–430. https://doi.org/10.1145/2818048.2819922

[44] Michael I Norton, Daniel Mochon, and Dan Ariely. 2012. The IKEA effect: When labor leads to love. *Journal of consumer psychology* 22, 3 (2012), 453–460.

[45] Catherine S. Oh, Jeremy N. Bailenson, and Gregory F. Welch. 2018. A Systematic Review of Social Presence: Definition, Antecedents, and Implications. *Frontiers in Robotics and AI* 5 (Oct. 2018), 114. https://doi.org/10.3389/frobt.2018.00114

[46] Kevin Pfeil, Pamela Wisniewski, and Joseph LaViola Jr. 2019. An Analysis of User Perception Regarding Body-Worn 360° Camera Placements and Heights. In *ACM Symposium on Applied Perception 2019* (Barcelona, Spain) *(SAP '19)*. ACM, New York, NY, USA, Article 13, 10 pages. https://doi.org/10.1145/3343036.3343120

[47] Kevin P. Pfeil, Karla Badillo-Urquiola, Jacob Belga, Jose-Valentin T. Sera-Josef, Joseph J. Laviola, and Pamela J. Wisniewski. 2024. Using Co-Design with Streamers and Viewers to Identify Values and Resolve Tensions in the Design of Interpersonal Wearable Telepresence Systems. 8, CSCW1, Article 148 (apr 2024), 21 pages. https://doi.org/10.1145/3637425

[48] Kevin P. Pfeil, Neeraj Chatlani, Joseph J. LaViola, and Pamela Wisniewski. 2021. Bridging the Socio-Technical Gaps in Body-Worn Interpersonal Live-Streaming Telepresence through a Critical Review of the Literature. *Proc. ACM Hum.-Comput. Interact.* 5, CSCW1, Article 120 (April 2021), 39 pages. https://doi.org/10.1145/3449194

[49] Jason Procyk, Carman Neustaedter, Carolyn Pang, Anthony Tang, and Tejinder K. Judge. 2014. Exploring video streaming in public settings: shared geocaching over distance using mobile video chat. In *Proceedings of the 32nd annual ACM conference on Human factors in computing systems - CHI '14*. ACM Press, Toronto, Ontario, Canada, 2163–2172. https://doi.org/10.1145/2556288.2557198

[50] Irene Rae, Gina Venolia, John C. Tang, and David Molnar. 2015. A Framework for Understanding and Designing Telepresence. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing - CSCW '15*. ACM Press, Vancouver, BC, Canada, 1552–1566. https://doi.org/10.1145/2675133.2675141

[51] Mel Slater, Martin Usoh, and Anthony Steed. 1994. Depth of presence in virtual environments. *Presence: Teleoperators & Virtual Environments* 3, 2 (1994), 130–144. Publisher: MIT Press.

[52] Andrew Sommerlad, Louise Marston, Jonathan Huntley, Gill Livingston, Gemma Lewis, Andrew Steptoe, and Daisy Fancourt. 2022. Social relationships and depression during the COVID-19 lockdown: longitudinal analysis of the COVID-19 Social Study. *Psychological medicine* 52, 15 (2022), 3381–3390.

[53] Daxton R. "Chip" Stewart and Jeremy Littau. 2016. Up, Periscope: Mobile Streaming Video Technologies, Privacy in Public, and the Right to Record. *Journalism & Mass Communication Quarterly* 93, 2 (June 2016), 312–331. https://doi.org/10.1177/1077699016637106

[54] Anthony Tang, Omid Fakourfar, Carman Neustaedter, and Scott Bateman. 2017. Collaboration with 360° Videochat: Challenges and Opportunities. In *Proceedings of the 2017 Conference on Designing Interactive Systems* (Edinburgh,

United Kingdom) *(DIS '17)*. Association for Computing Machinery, New York, NY, USA, 1327–1339. https://doi.org/10.1145/3064663.3064707

[55] John Tang, Gina Venolia, Kori Inkpen, Charles Parker, Robert Gruen, and Alicia Pelton. 2017. Crowdcasting: Remotely Participating in Live Events Through Multiple Live Streams. *Proceedings of the ACM on Human-Computer Interaction* 1, CSCW (Dec. 2017), 1–18. https://doi.org/10.1145/3134733

[56] Hiroaki Tobita. 2017. Gutsy-Avatar: Computational Assimilation for Advanced Communication and Collaboration. In *2017 First IEEE International Conference on Robotic Computing (IRC)*. IEEE, Taichung, Taiwan, 8–13. https://doi.org/10.1109/IRC.2017.82

[57] Bob G Witmer and Michael J Singer. 1998. Measuring presence in virtual environments: A presence questionnaire. *Presence* 7, 3 (1998), 225–240. Publisher: MIT Press.

[58] Yanlai Wu, Xinning Gui, Pamela J. Wisniewski, and Yao Li. 2023. Do Streamers Care about Bystanders' Privacy? An Examination of Live Streamers' Considerations and Strategies for Bystanders' Privacy Management. *Proc. ACM Hum.-Comput. Interact.* 7, CSCW1, Article 127 (apr 2023), 29 pages. https://doi.org/10.1145/3579603

[59] Svetlana Yarosh, Kori M. Inkpen, and A.J. Bernheim Brush. 2010. Video Playdate: Toward Free Play Across Distance. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Atlanta, Georgia, USA) *(CHI '10)*. ACM, New York, NY, USA, 1251–1260. https://doi.org/10.1145/1753326.1753514

[60] Takeru Yazaki, Daisuke Uriu, Yuna Watanabe, Ryoko Takagi, Zendai Kashino, and Masahiko Inami. 2023. "Oh, Could You Also Grab That?": A Case Study on Enabling Elderly Person to Remotely Explore a Supermarket Using a Wearable Telepresence System. In *Proceedings of the 22nd International Conference on Mobile and Ubiquitous Multimedia* (<conf-loc>, <city>Vienna</city>, <country>Austria</country>, </conf-loc>) *(MUM '23)*. Association for Computing Machinery, New York, NY, USA, 340–352. https://doi.org/10.1145/3626705.3627799

[61] Takeru Yazaki, Yuna Watanabe, Lingrong Kong, and Masahiko Inami. 2023. Design and Field Study of Syn-Leap: A Symmetric Telepresence System for Immersion Switching and Walking Across Multiple Locations. In *Proceedings of the 22nd International Conference on Mobile and Ubiquitous Multimedia* (<conf-loc>, <city>Vienna</city>, <country>Austria</country>, </conf-loc>) *(MUM '23)*. Association for Computing Machinery, New York, NY, USA, 353–365. https://doi.org/10.1145/3626705.3627772

[62] Jacob Young, Tobias Langlotz, Matthew Cook, Steven Mills, and Holger Regenbrecht. 2019. Immersive Telepresence and Remote Collaboration using Mobile and Wearable Devices. *IEEE Transactions on Visualization and Computer Graphics* 25, 5 (May 2019), 1908–1918. https://doi.org/10.1109/TVCG.2019.2898737

[63] Jacob Young, Tobias Langlotz, Steven Mills, and Holger Regenbrecht. 2020. Mobileportation: Nomadic Telepresence for Mobile Devices. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 4, 2 (June 2020), 1–16. https://doi.org/10.1145/3397331

Received January 2024; revised April 2024; accepted May 2024