RESEARCH-ARTICLE

Shared Realities: Avatar Identification and Privacy Concerns in Reconstructed Experiences

CHENGYAO WANG, Cornell University, Ithaca, NY, United States

SANDHYA SRIRAM, Cornell University, Ithaca, NY, United States

ANDREA STEVENSON WON, Cornell University, Ithaca, NY, United States

Proceedings of the ACM on Human-Computer Interaction, Volume 5, Issue CSCW2 (October 2021)
https://doi.org/10.1145/3476078
EISSN: 2573-0142

Shared Realities: Avatar Identification and Privacy Concerns in Reconstructed Experiences

CHENG YAO WANG, Cornell University, USA
SANDHYA SRIRAM, Cornell University, USA
ANDREA STEVENSON WON, Cornell University, USA

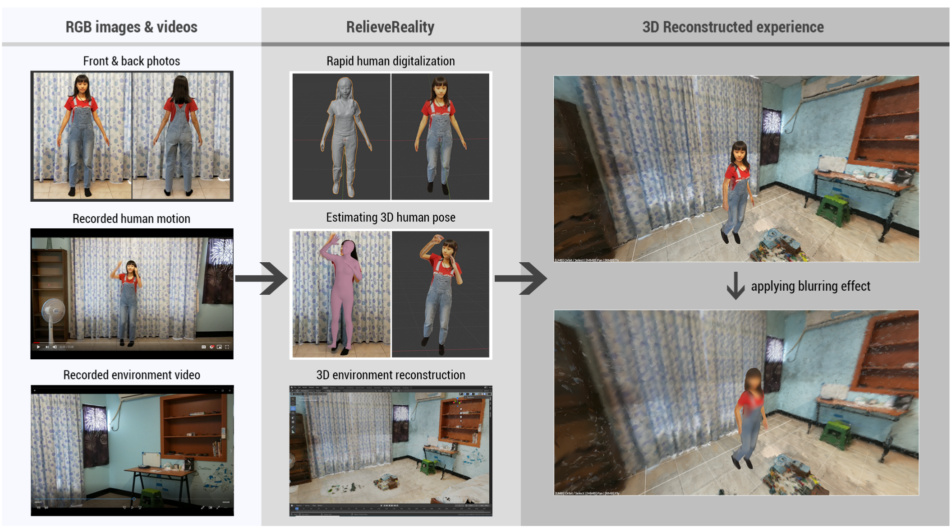

Fig. 1. ReliveReality allows people to reconstruct experiences in 3D from RGB videos and share them as a new experience-sharing method.

Recent advances in 3D reconstruction technology allow people to capture and share their experiences in 3D. However, little is known about people's sharing preferences and privacy concerns for these reconstructed experiences. To fill this gap, we first present ReliveReality, an experience-sharing method utilizing deep learning-based computer vision techniques to reconstruct clothed humans and 3D environments and estimate 3D pose with only a RGB camera. ReliveReality can be integrated into social virtual environments, allowing others to socially relive a shared experience by moving around the experience from different perspectives, on desktop or in VR. We conducted a 44-participant within-subject study to compare ReliveReality to viewing recorded videos, and to a ReliveReality version with blurring obfuscation. Our results shed light on how people identify with reconstructed avatars, how obfuscation affects reliving experiences, and sharing preferences and privacy concerns for reconstructed experiences. We propose design implications for addressing these issues.

CCS Concepts: • **Human-centered computing** → **Virtual reality**; *Social content sharing*.

Authors' addresses: Cheng Yao Wang, cw776@cornell.edu, Cornell University, USA, Ithaca, NY; Sandhya Sriram, ss3769@cornell.edu, Cornell University, USA, Ithaca, NY; Andrea Stevenson Won, Cornell University, USA, Ithaca, NY.

2573-0142/2021/October-ART337 $15.00
https://doi.org/10.1145/3476078

Additional Key Words and Phrases: 3D reconstructed experience, privacy sharing digital contents, privacy-enhancing techniques, obfuscation

ACM Reference Format:
Cheng Yao Wang, Sandhya Sriram, and Andrea Stevenson Won. 2021. Shared Realities: Avatar Identification and Privacy Concerns in Reconstructed Experiences . *Proc. ACM Hum.-Comput. Interact.* 5, CSCW2, Article 337 (October 2021), 25 pages. https://doi.org/10.1145/3476078

1 INTRODUCTION

People often share their experiences through photos and videos to entertain others, strengthen social connections, or simply for the self-fulfillment of sharing [74]. However, photos and videos only allow people to relive experiences from the perspective of the camera. To overcome this limitation, we present the ReliveReality prototype, which enables an experience-sharing method that allows people to reconstruct their experiences in 3D and share them with others in social virtual environments. As shown in Figure 1, by leveraging deep learning-based computer vision techniques, ReliveReality requires only a single RGB camera for the full pipeline of reconstructing experiences in 3D. With this, we generate self-similar avatars from users' front and back photos, and estimate 3D human poses and reconstruct 3D environments from RGB videos.

ReliveReality allows people to 'enter' into others' past experiences, move around in these environments, and relive events from different perspectives. Prior work [79] has shown that this freedom to explore relived experiences allows users to understand shared experiences better and discover previously unnoticed things, making the sharing experience more fulfilling. However, despite the many benefits, sharing highly detailed reconstructed experiences also raises privacy concerns and other critical yet unexplored research questions.

First of all, at the heart of the reconstructed experience is a reconstructed avatar resembling the user and performing the user's recorded actions. To better understand the user experience when people share their reconstructed experiences, we examine *RQ1: How do people identify themselves with reconstructed avatars that look like them?*

Second, sharing reconstructed experiences may raise privacy concerns since both reconstructed avatars and reconstructed environments are rich with information about people's identities, personalities, values, and lifestyles. In addition, people can interact with the reconstructed avatars of others in ways they may not approve, or move closer to examine details of reconstructed environments that sharers did not intend to share. Therefore, it's important to investigate *RQ2: What are people's sharing preferences and privacy concerns for reconstructed experiences?*

Lastly, previous privacy research on photos and videos [31, 46] has proposed various privacy-enhancing obfuscation methods such as inpainting and blurring. However, the effectiveness of applying these obfuscations to a reconstructed experience is unclear, leading to *RQ3: How does obfuscating reconstructed experiences affect users' avatar identification and privacy concerns?*

We conducted a two-stage 1x3 within-subject study with 44 participants. In Part 1, participants mimicked three funny dances and recorded their experience with (1) front and back photos of themselves, (2) videos of their dance movements, and (3) a video capturing their surrounding environments. We then utilized ReliveReality to generate reconstructed experiences, which included participants' avatars, their dance movements, and their environments. In Part 2, participants relived their dance experiences in three conditions. These included 1) viewing a 2D recorded video (ViewVideo), 2) reliving a reconstructed experiences (ReliveReality), and 3) reliving a reconstructed experiences with blurring obfuscation (BlurReality). We compared avatar identification, presence, and participants' sharing preferences and privacy concerns in each condition, along with a semi-structured interview to elicit participants' qualitative responses.

Our results suggest that participants prefer to share highly detailed reconstructed experiences with people who they know well (i.e. strong ties [37]), but with their weak ties (e.g. distant friends or relatives, friends of friends) [29], they prefer sharing obfuscated, blurred versions to reduce privacy concerns. The contributions of this paper are fourfold:

- Presenting the ReliveReality prototype which allows people to reconstruct their experiences in 3D with only a RGB camera and share them in social virtual environments.
- Applying the behavioral privacy model to investigate people's preferences and privacy concerns in sharing personal reconstructed experiences.
- Understanding how people identify with their reconstructed avatars, their privacy concerns, and how obfuscation affects the reconstructed experience through a 44-participant within-subject mixed-methods study.
- Proposing design implications for sharing reconstructed experiences.

2 RELATED WORK

Our work combines and extends earlier research in reconstructing clothed humans, poses and environments, avatar identification with self-similar avatars, and preserving privacy in digital content. In this section, we review key papers from earlier work to contextualize our novel research contribution.

2.1 Reconstructing clothed humans, poses and environments

2.1.1 Rapid human digitalization. 3D human reconstruction has been explored for several decades in the field of computer vision and computer graphics. Accurate methods based on stereo or fusion have been proposed using various types of sensors [11, 54, 65, 71], and several applications have become popular in sports, medicine and entertainment (e.g., movies, games, AR/VR experiences). However, these setups require multi-camera rigs or RGB-D cameras. Recently, rapid progress in deep neural networks has made full 3D human reconstruction with detailed geometry and appearance possible, whether using a few images [33, 67, 68] or a video sequence [4, 5, 61]. In our ReliveReality prototype, we utilized PIFuHD [68], a multi-level framework that performs joint reasoning over holistic information and local details and uses image-to-image translation networks to handle uncertainty in regions that are not observed (e.g. the back) to achieve high-resolution 3D reconstructions of clothed humans from a single image.

2.1.2 3D Human pose estimation from videos. Motion capture from monocular video offers many advantages, such as simple setup, low cost, and a non-intrusive capture process. While 3D human pose estimation from videos is highly challenging due to depth ambiguities and occlusions, significant progress has been achieved in recent years by data-driven learning-based approaches. Due to the availability of video datasets associated with ground truth motion capture data (e.g., Human3.6M [35] and HumanEva [72]), some methods (Mehta et al. [50, 51]) trained end-to-end models to extract the poses directly from the image pixels. However, they do not perform well on in-the wild datasets like 3DPW [78] and MPI-INF-3DHP [20]. Another family of 3D pose estimators [20, 59, 63] uses a two-stage approach to "lift" off-the-shelf 2D keypoints [15, 16, 23] into 3D joint locations either by regression [17, 49] or model fitting [12]. A parallel line of research aims at jointly recovering human shape and pose [39, 42, 81]. In our ReliveReality prototype, we use VIBE [42], which trained a temporal model to predict the parameters of the SMPL body model for each frame while a motion discriminator tried to distinguish between real and regressed sequences. It achieves state-of-the-art performance on the in-the wild 3DPW dataset and have been utilized for reconstructing 3D human poses from online videos in other recent work [27, 30].

2.1.3 3D environment reconstruction. In recent years, the three-dimensional (3D) reconstruction of semantically rich and geometrically accurate indoor environments has emerged as a significant and challenging task. Most of the current indoor 3D models acquisition technologies are based on LiDAR [22], Kinect depth cameras [55], or image-based approaches such as robot simultaneous localization and mapping (SLAM) [2] or photogrammetry [52]. A comprehensive review of all state-of-the-art techniques for the 3D reconstruction of indoor environments is outside the scope of our paper but has been well reported in [40]. Since our goal is to reconstruct users' environments with only RGB videos, we leveraged the photogrammetry approach provided by Meshroom [1], an open source photogrammetric computer vision framework.

To our best knowledge, ReliveReality is the first prototype that integrates various deep-learning based computer-vision techniques to enable people to reconstruct their experiences in 3D, including generating detailed avatars of clothed humans, estimating 3D human poses and reconstructing the 3D environment with only a single RGB camera. In addition, future advanced avatar creation and pose estimation methods can be easily plugged into the ReliveReality framework, which makes the proposed framework flexible and generalizable.

2.2 Avatar identification with self-similar avatars and doppelgangers

The multi-faceted concept of avatar identification includes the degree to which the avatar is similar to the user [18, 75, 76]. Prior research has shown that identifying with an avatar has positive outcomes for play experience and enjoyment [75], shapes our behaviour outside of the game [82], and makes us more susceptible to persuasive messages [53].

Both customization and self-similarity have been shown to influence identification. For example, Trepte et al. [75] showed that customizing an avatar increases identification and leads to higher enjoyment, as does creating an avatar that has high similarity to the player. Freeman et. al. [25] conducted an in-depth analysis of the various factors that users might consider for choosing their avatars in a social virtual environment. Their results suggested that users tend to construct self-presentations that are similar to their physical selves. Koulouris et. al. [43] investigated the effects of avatar customisation and identification on the player experience and physical performance in a high-intensity VR exergame. Their results showed that customizing avatars by selecting physical features using the avatar creation tool can result in greater similarity identification, wishful identification, importance to identity and greater emotional investment, compared to a generic avatar. Gonzalez-Franco et. al. [28] also found that self-identification on avatars can be increased through adding pre-baked facial animation. The benefits of self-similar avatars also extends to education; in a study by Parmar, students learned programming concepts better in a virtual reality learning environment when they used self-similar, embodied avatars [58].

There is also a growing body of research related to the psychological effects of viewing a self-similar agent-avatar (a virtual "doppelganger"). In health research, Fox and Bailenson [24] found that participants who saw their own avatar lose or gain weight based on physical activity levels engaged in more voluntary exercise during the study than those whose avatar did not change or who witnessed weight change on a different avatar's body. In addition, the self-observing participants also reported exercising more in the 24 hour period following the study [24]. Advertisements with doppelgängers as compared to to representations of others tend to be more effective [3]. However, watching at a virtual double is not always beneficial and can be harmful. In preparation for public speaking, participants who watched non-similar characters speak reduced anxiety compared to watching virtual doubles giving the talk [7] and Helen et al. [80] found that there was no effect of avatar appearance on players' performance or subjective experience in a search and rescue game. Junuzovic et al. [36] also showed that participants reported self-similar cartoon avatars to be inappropriate for a professional discussion.

Unlike the reconstructed avatars in our study, the avatars in the above studies were not photorealistic. Recent advances in computer vision and 3D scanning technology have enabled the rapid creation of highly photorealistic, self-similar 3D avatars from human subjects using image, video or depth sensing cameras. However, most of the existing research in this area focuses on the process of generating highly photorealistic avatars and only a few studies investigate avatar identification with photorealistic self-similar avatars. (For example, Helen et al. [80] measured avatar identification in a search-and-rescue computer game when using photorealistic avatars of oneself or of a friend. They found that participants identified significantly more with their own avatar than their friend's).

To our knowledge, no previous research has examined the effects of a highly realistic self-similar avatar that *was*, but *is no longer*, controlled by the self, on identification and privacy concerns.

2.3 Preserving privacy in sharing digital contents

2.3.1 Privacy theories and framework. Privacy is a multifaceted concept and researchers from different disciplines, including philosophy, law, communication, social psychology, and Human-Computer Interaction (HCI) have proposed various theories, such as Communication Privacy Management Theory (CPM) [60], Contextual Integrity (CI) [56] and Behavioral Theory of Privacy [14] to study privacy from various perspectives. Similar to photos and videos, 3D reconstructed experiences that contain rich visual information (e.g. photorealistic avatars and environments) are likely to cause privacy issues when shared online. To ground our study design within a privacy theory, we looked into the behavioral theory of privacy [14], which has been shown to be particularly useful in the context of sharing photos [45, 46]. The behavioral theory of privacy focuses on two key elements that influence privacy – information content, and information recipient. Adjusting either recipient or content would affect privacy concerns. Focusing on these two elements help us to answer our main research questions about how reconstructed experiences affect people's sharing preference and privacy concerns. In addition, the behavioral theory of privacy also examines users' privacy enhancing behaviors and clusters them into three temporal categories: avoidance, which occurs prior to an act; modification, which occurs during an act; and alleviation, which occurs after an act. Investigating privacy from a temporal perspective has been found useful in understanding how users manage privacy in a collaborative environment [62]. Observing at what point in time certain privacy behaviors are enacted can broaden our understanding of sharing experiences in 3D virtual environments. These temporal aspects of privacy also helped us organize our findings and made our design implications more generalizable.

2.3.2 Privacy enhancing techniques for photos and videos. Sharing personal digital contents such as photos and videos provides a natural mechanism for people to express themselves and interact with one another [38]. With the dramatic increase of online sharing practice, prior work on reducing privacy risks in photos and videos mainly falls into two broad categories [46]: controlling access, and limiting information content. Access control mechanisms allow people to specify who can access their shared content. However, managing access control settings requires substantial time and effort [47] and photos taken in public places may pose privacy threats to people who are not subjects of the photo. In addition, these methods might weaken one of the primary motivations for using social networking platforms: reaching out and making new connections [57]. On the other hand, controlling photo/video content, such as obscuring parts of content to protect privacy, has been widely studied and adopted by many existing applications. In the context of remote collaboration via live video feed, Boyle et al. [13] studied how blurring and pixelating affect privacy and awareness, and Hudson et al. [34] proposed techniques such as representing people's movement using dark pixels overlaid on a static image to reduce privacy risks while keeping information

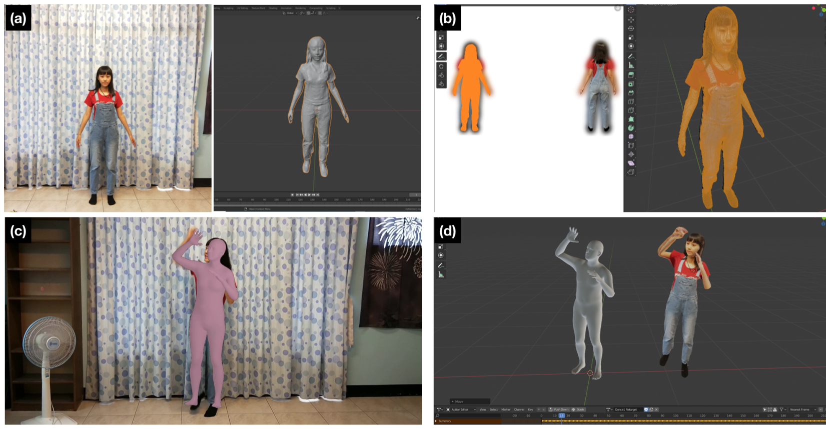

Fig. 2. (a) Rapid human digitalization via PIFuHD [68]. (b) Texturing front part of the avatar in Blender. (c) 3D multi-person human pose estimation via VIBE [42]. (d) Retargeting 3D pose from the SMPL model to the reconstructed avatar.

required for the collaboration. To reduce the privacy risks of online photo sharing, recent studies have proposed using privacy-enhancing image obfuscations to obscure sensitive regions of photos [31], while trying to preserve the viewer's experience. These studies identify a set of obfuscations that can effectively obscure objects (or their properties) in a photo while minimizing the impact on the viewer's overall satisfaction. However, to the best of our knowledge, none of these privacy-enhancing techniques have been studied and applied in the context of sharing 3D reconstructed experiences. In this paper, we grounded our study design in behavioral privacy theory [14] and examined how different recipients affected people's sharing preference and privacy concerns. We also consider for the first time how obfuscation affects identification with reconstructed avatars and the resulting reliving experience.

3 RELIVEREALITY PROTOTYPE

3.1 Design Goals

Our main design goals for the ReliveReality experience were that it be accessible, easy to use, generalizable and easy to share.

3.1.1 Accessible. ReliveReality allows users to simply reconstruct their experience via a single RGB camera without the need for other 3D scanning devices or professional multi-camera rigs. Making the process reliant on a single RGB camera not only makes ReliveReality more accessible but also enables researchers to conduct 3D reconstruction-related research via other experimental methods beyond lab experiments. For example, remote experiments were particularly useful during the COVID-19 pandemic.

3.1.2 Ease of Use. We focused on automating and simplifying the three most complicated steps in the full reconstruction pipeline including reconstructing avatars, poses, and environments. We leveraged existing deep-learning based computer vision techniques and implemented them directly on Google Colab. Thus, users can directly run our prototype on their browsers without the need for powerful computers and setting up deep learning environments. Users can reconstruct their avatars, poses, and environment by simply uploading their photos and video to Google Colab and run the prototype. However, we note that users still need to manually complete other steps such as texturing the models, applying motions to avatars, and uploading to Mozilla Hubs. Although it's possible to realize a full automatic reconstruction pipeline, it's beyond the main focus of the current paper, which primarily investigates privacy concerns in the context of reconstructed avatars and environments. We make the code publicly available so other researchers and users can further develop these techniques.

3.1.3 Generalizable. As 3D reconstruction technology continues to evolve, it's important to make the ReliveReality framework generalizable and extendable so that the modules in the framework can be replaced as advances are made. We achieved this by dividing the whole reconstructing experience into independent modules.

3.1.4 Easy to share. Allowing users to easily share their 3D reconstructed experience is critical for investigating their sharing preferences and privacy concerns. To achieve this goal, we utilized WebGL techniques and the Mozilla Hubs social virtual environment to enable sharing reconstructed experience through website links.

3.2 ReliveReality Framework

The ReliveReality framework (Figure 1) of capturing and reconstructing experiences in 3D consists of five modules: (1) Rapid human digitalization: generating detailed 3D human avatars from users' front and back photos. (2) Texturing and rigging: producing animatable 3D rigged full-body human avatars with color textures. (3) 3D multi-person pose estimation: estimating the 3D pose for each detected human in videos. (4) Motion-retargeting: retargeting estimated 3D motions to animate reconstructed avatars. (5) 3D environment reconstruction: recreating the environment from videos.

3.2.1 Rapid human digitalization. For rapidly digitizing detailed clothed humans, we utilized PIFuHD [68], which offers a multi-level framework to infer the 3D geometry of clothed humans with high detail from a single photo of the user, as shown in Figure 2 (a).

3.2.2 Texture and Rigging. As shown in Figure 2 (b), we utilized the 'Project from View' technique in Blender from both front and back views and manually aligned each unwrapped UV to the user's front and back photos respectively. Edge blurring is applied to make the texture more robust to error-prone manual alignment processes. After texturing, we rigged the reconstructed avatars with Adobe Mixamo's online auto-rigger.

3.2.3 3D multi-person human pose estimation. We utilized a deep-learning based approach open sourced as VIBE [42], which trains a temporal model to predict the parameters of the SMPL body model [48] for each frame while a motion discriminator tries to distinguish between real and regressed sequences as shown in Figure 2(c). We extracted 3D joint rotations data from the predicted SMPL pose parameter and saved it for animating reconstructed avatars.

3.2.4 Motion Retargeting. Since the rest pose and original joint rotations of the SMPL skeleton and the reconstructed avatar are different, we implemented a custom script to retarget the predicted 3D pose of the SMPL model to the reconstructed avatar as shown in Figure 2(d).

3.2.5 3D environment reconstruction. We used MeshRoom[1] (Figure 3(a)), an open source photogrammetric computer vision framework, to reconstruct 3D environments from RGB videos.

3.2.6 Sharing reconstructed experiences via social virtual environments. We utilized the Mozilla Hubs social virtual environment (SVE) to enable users to share their reconstructed experiences with others. Figure 3 (b)(c)(d) demonstrates some Hub rooms showing participants' animated avatars and reconstructed environments. The Hub rooms are private and only accessible through the link

Fig. 3. (a) 3D environment reconstruction via Meshroom (b)-(d) Hub rooms with reconstructed avatars and environments.

we generated. Users can move around in the reconstructed experience via first-person game-like mechanics, using their keyboard and mouse to move on 2D displays, or teleporting in VR.

4 USER STUDY

We conducted a two-stage 1x3 within-subject study with 44 participants to compare avatar identification, presence and sharing preferences in three conditions: (1) viewing recorded video (ViewVideo), (2) reliving reconstructed experience (ReliveReality), and (3) reliving reconstructed experience with blur obfuscation (BlurReality), as shown in Figure 4. We selected blurring as the obfuscation method in our study as it is the most commonly used method to control disclosure both in research and in practice. We chose a within-subjects design for its greater statistical power and also to better understand the differences between the conditions through participants' qualitative responses. To address issues such as carryover effects and expectancy bias, we counterbalanced the orders of conditions. For 44 participants, each of the 6 possible permutations of 3 study conditions appeared 7 times in random order. An extra 2 random permutations were assigned to last 2 participants. In addition, we also randomized the funny dance movements used in each condition.

4.1 Participants

A total of 44 participants (14 m, 30 f) were recruited. All participants were recruited from the undergraduate student population of a medium-sized private university. Participants were aged 18-24 ($M = 21.5, SD = 3.0$). Both the first part and the second part of the study took around 1 hour to complete and participants were compensated with \$10 gift cards in each study. IRB approved all aspects of the experiment, and all participants signed informed consent. A subset of participants consented separately to having their images shared publicly.

4.2 Mimicking funny dances as the reconstructed experience

In order to let participants have an engaging experience that was worth recording and reliving in our study, participants were asked to view funny dance videos and then mimic the dance movements. The task of mimicking funny dances was chosen because it contains various human motions that make the reconstructed experience more interesting and worth sharing. We had

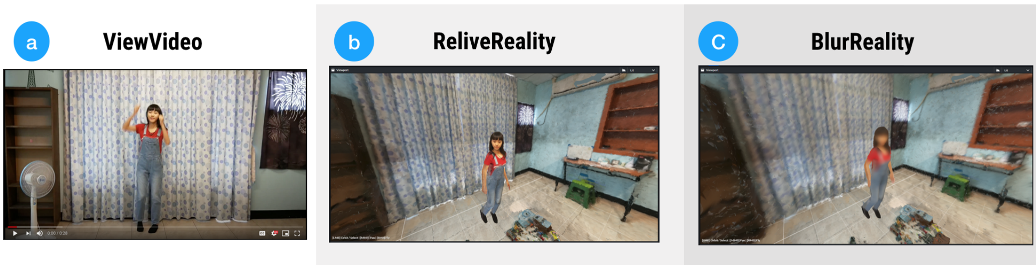

Fig. 4. Three methods for reliving experiences: (a) ViewVideo: viewing recorded video (ViewVideo);(b) ReliveReality: reliving reconstructed experience; (v) BlurReality: eliving reconstructed experience with blurring obfuscation.

conducted a pilot study with 8 participants, who reported having fun and feeling comfortable while mimicking the selected funny dances. More importantly, we found that participants reported different sharing preferences when sharing their recorded funny dance experiences to different recipient groups. While this task necessarily has limitations, we felt it was an appropriate starting point to systematically investigate sharing preferences and privacy concerns.

4.3 Procedure

Due to the current COVID-19 pandemic, we conducted the study entirely online. Researchers gave instructions using the Google Meet video conferencing tool.

4.3.1 First stage: Funny dance experiences and preparing reconstruction materials. In the first stage, participants were guided through collecting the reconstruction materials. These included (1) participants' front and back photos, (2) recording three 30-second videos of themselves dancing, with their full body visible, and (3) a video of their environment. To create the videos, participants viewed three 30-second "funny dance" videos and then mimicked each dance for 30 seconds.

4.3.2 Generating reconstructed experience for all participants. After getting the participants' materials from the first part of the study, we utilized the ReliveReality prototype to generate reconstructions for the ReliveReality condition. We then applied gaussian blur to the textures of the reconstructed avatars and environments to generate the BlurReality reconstruction shown in Figure 1.

4.3.3 Second stage: Reliving reconstructed experiences. After one week, all participants who completed the first part of the study returned for the second part of the study. Participants were asked to relive their three dance experiences in each condition (1) viewing their recorded dance video (ViewVideo), (2) reliving the reconstructed experience (ReliveReality) and (3) reliving the reconstructed experience with blur obfuscation (BlurReality). In the ViewVideo condition, participants were asked to watch their recorded dance video. For the ReliveReality and BlurReality conditions, they were asked to view the reconstructed experiences in Mozilla Hubs. Only one of three funny dances were relived in each condition and the order of the condition and the funny dance shown in each were randomized. After each condition, participants completed a questionnaire. Finally, a semi-structured interview was conducted with each participant to elicit their qualitative responses and better understand the reasons behind their answers in the questionnaire.

4.4 Measures and Questionnaire

Our questionnaire measured participants' social media usage, avatar identification, presence, and their sharing preferences for each study condition. We also used a survey tool developed by

Order	Module	Auto/Manual	Platform	Time (unit)
1	Rapid human digitalization	Auto	Google Colab	4.26 (sec)
2	Texturing & rigging	Manual	Blender & Adobe Mixamo	~10 (mins)
3	3D human pose estimation (3 x 30-sec single-personvideos)	Auto	Google Colab	2. 31 (mins)
4	Motion retargeting & Animation generations	Auto	Blender & Custom scripts	2.66 (mins)
5	3D environment reconstruction (avg. 259 photos)	Auto	Google Colab	77.84 (mins)
6	Integrating into Mozilla Hub	Manual	Hub rooms editor (i.e Spoke)	~ 10 (mins)

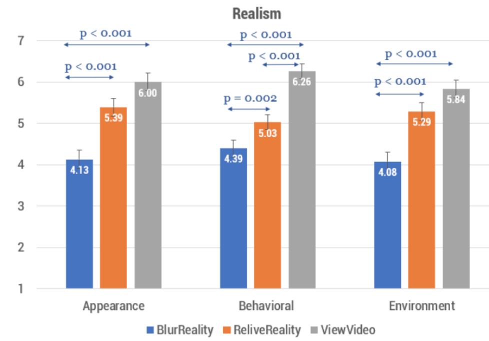

Fig. 5. (Left) The performance of each module in the ReliveReality prototype. (Right) The quality of the recontructed experience including the appearance and behavioral realism of avatars and realism for environments.

Vastenburg et al [77] to capture multiple emotion categories with an easy-to-use pictorial mood-reporting instrument in each condition. In addition, we designed interview questions to discuss the overall experience of the study, participants' preferences and their judgment of the quality of each reconstructed experience along with its nuances and challenges. As in other prior work [69, 70], questions for measuring presence were adapted from the igroup presence questionnaire (IPQ) [64].

4.4.1 Avatar Identification. To generate the questions for measuring avatar identification, we combined validated subscales suitable for use in the context of reconstructed avatars from the aforementioned Player-Identification Scale (PIS) [76], Player-Avatar Identification Scale (PAIS) [44] and Player-Avatar Interaction Scale (PAX) [9]. Table 1 shows the resulting new questionnaire.

4.4.2 Comfort with Sharing. Participants rated the comfortability of sharing reconstructed content with each of nine recipient groups including significant others, close family members, close friends, close colleagues/classmates, distant family members, distant friends, distant colleagues/classmates, friends of friends and people only met online. These recipients were developed based on prior work [14, 45] that also applied the behavioral theory of privacy to investigate sensitive contents and privacy in online photo sharing. Participants answered "How comfortable are you if you are going to share this with [recipient groups]" on a Likert-type scale from 1-"very uncomfortable" to 7-"very comfortable" for each condition.

5 RESULTS

A similar statistical approach was used for all data reported in this section. We ran linear mixed model analysis in R using the lme4 package. To account for the random effects that arise from the individual participant, we included participants as a random effect. The different conditions were dummy-coded and treated as a fixed factor (i.e. three types of reliving content) with three levels (i.e., ViewVideo, ReliveReality and BlurReality). We applied ANOVAs to each linear mixed model. Significance was calculated using the lmerTest package, which applies Satterthwaite's method to estimate degrees of freedom and generate p-values for mixed models. We further examined our data with Tukey's pairwise comparison and Tukey adjustments was done by obtaining estimated marginal means (i.e. EMMs).

5.1 Performance and quality of reconstructed experience

5.1.1 Performance. Figure 5 (left) summarizes the required time for each module of the ReliveReality. Overall it took around 25 minutes for all steps except for the 3D environment reconstruction. Generating the reconstructed environment from an average of 295 photos per environment requires 1.5 hours on average. Photogrammetric-based reconstruction will fail under some environmental situations such as reflective surfaces, plain walls and poor lighting conditions. Thus, 6 out of 44 participants were excluded from the second part of the study due to failed reconstructions. In the end, we prepared 38 unique reconstructed experiences for the second part of the study.

5.1.2 Realism. To better understand the quality of the reconstructed experience, we collected subjective ratings for different aspects of realism, including avatar's appearance realism, avatar's behavior realism and environment's realism, for each reliving content. The means and standard deviations for avatar appearance realism ratings are: ViewVideo=6(1.32), ReliveReality=5.39(1.26), BlurReality=4.13(1.40). In terms of avatar's behavioral realism, the means and standard deviations are: ViewVideo=6, .26(1.08), ReliveReality=5.03(1.17), BlurReality=4.39(1.28). For the environment realism, the means and standard deviations are: ViewVideo=5.84(1.31), ReliveReality=5.29(1.35), BlurReality=4.08(1.42). In addition, we created three linear mixed models to analyze and compare each realism among 3 different types of reliving contents. As shown in Figure 5 (Right). ANOVA for all linear mixed models yielded significant effects of the types of reliving contents (Appearance Realism: $F(2, 74) = 23.04, p < .0001$; Behavior Realism: $F(2, 74) = 36.98, p < 0.0001$; Environment Realism: $F(2, 74) = 21.00, p < 0.0001$). Pairwise comparison of avatar appearance realism and environment realism both showed significant differences between ViewVideo and BlurReality (Appearance Realism: $p < 0.0001, t = 6.65$; Environment Realism: $p < 0.0001, t = 6.34$) and between BlurReality and ReliveReality (Appearance Realism: $p = 0.0001, t = -4.50$; Environment Realism: $p = 0.0001, t = -4.35$), but not between ViewVideo and ReliveReality ($p > 0.05$). In terms of avatar behavior realism, pairwise comparison showed significant differences between BlurReality ($p < 0.0001, t = 8.45$; Environment Realism: $p < 0.0001, t = 6.34$), between BlurReality and ReliveReality ($p = 0.002, t = -2.86$), and between ViewVideo and ReliveReality ($p < 0.0001, t = -5.60$). These results suggest that the ReliveReality prototype can generate avatars resembling participants and reconstruct the 3D environment with high fidelity, but the movements of the avatar are less realistic than the ones in the video.

5.2 Factor Analysis for Questionnaire Data

We ran an exploratory factor analysis (EFA) [19] to better understand the important factors in our questions for measuring avatar identification and presence, which were adapted from several prior work as mentioned above. Bartlett's test of sphericity was significant ($\chi 2(2, 300) = 2548.779, p < 0.001$) indicating that it was appropriate to use the factor analytic model on this set of data. The Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy showed that the strength of the relationships among variables was greater than 0.5 (KMO=0.88), thus it was acceptable to proceed with the analysis. Both an analysis of the scree plot and Kaiser's criterion suggested a three-factor structure. Furthermore, since we assumed that factors would be related, we used oblique rotation (”oblimin”) along with standard principal axes factoring. To ensure the factors were meaningful, we only took items with factor loadings of 0.3 and above. As shown in Table 1, the only excluded question was Q6- “How aware were you of the real world surrounding while reliving the experience?” One plausible reason is that this question was originally designed and used in the context of immersive VR experiences which do not fit our study in which participants viewed the 3D reconstructed experiences on 2D displays. The cumulative explained variance of the three factors is 62.2%. The 22 questionnaire items in bold were used for our evaluation of the three conditions (ViewVideo, ReliveReality, and BlurReality). The factor analysis result yielded 3 factors: *Physical Identification (PhI)*, *psychosocial Identification (PsyI)*, and *Presence and Satisfaction (PS)*. Informed by the avatar identificaiton model [41], we conceptualized the F1 and F2 factors to the physical identification and psychosocial identification. Physical identification deals with the level of anthropomorphism as well as the level of physical similarity between an avatar and a participant. This means that the more realistically human-looking and the more physically similar to an individual an avatar is, the greater physical identification s/he will have with that avatar. Psychosocial identification is a subjective measure of emotional and social connection with avatars. Cronbach's alpha was calculated on each factor and demonstrated high internal reliability with respective alphas of 0.96, 0.84, and 0.93.

5.3 Questionnaire Response Analysis

We consider the effects of the reliving conditions (ViewVideo, ReliveReality, BlurReality) on each factor with a 7-point likert-scale measure: *Physical identification*, *Psychosocial identification*, and *Presence and Satisfaction*. The results are shown in Figure 6 (a)(b)(c). The horizontal lines within each box represent the median, the box bounds the Inter-quartile (IQR) range, and the whiskers show the max and min non-outliers.

5.3.1 Physical Identification. The means and standard deviations for the Physical Identification questions (9 items) for each reliving condition are: ViewVideo=$5.94(1.02)$, ReliveReality=$5.06(1.17)$, BlurReality=$4.07(1.31)$. As shown in the Figure 6 (a), the ANOVA for the linear mixed model of physical identification yielded a significant effect of the reliving condition, $F(2, 74) = 40.54, p < .0001$. Pairwise contrasts showed significant differences between ViewVideo and BlurReality ($p < 0.0001, t = 9.00$), between ViewVideo and ReliveReality($p = 0.0002, t = 4.21$),and between ReliveReality and BlurReality ($p < 0.0001, t = -4.79$). These results indicate that ViewVideo was perceived to have the highest physical identification among the three conditions and the bluring effect in BlurReality sigificantly decreased the physical identification compared with ReliveReality condition.

5.3.2 Psychosocial Identification. The means and standard deviations for the Psychosocial Identification questions (7 items) for each reliving condition are: ViewVideo=$4.62(1.35)$, ReliveReality=$4.12(1.49)$, BlurReality=$3.58(1.52)$. Figure 6(b) shows a significant effect of the reliving condition, $F(2, 74) = 14.69, p < .0001$. Pairwise contrasts showed significant differences between ViewVideo and BlurReality ($p < 0.0001, t = 5.42$), between ViewVideo and ReliveReality($p = 0.03, t = 2.62$), and between ReliveReality and BlurReality ($p < 0.02, t = -2.79$). These results suggest that ViewVideo and ReliveReality both provided higher levels of psychosocial identification than BlurReality, and ViewVideo provided the highest levels of psychosocial identification.

5.3.3 Presence and Satisfaction. While multiple definitions of presence have been proposed [66], the presence in this study refers to the observer's sense of psychologically leaving their real location and feeling as if transported to a virtual environment. In other words, our presence definiton is more close to what Heeter describes as the illusion of "being there" [32]. The means and standard deviations for the Presence questions (6 items) for each sharing condition are: ViewVideo=$3.76(0.91)$, ReliveReality=$5.00(1.00)$, BlurReality=$4.31(0.92)$. As shown in the 6 (c), the the ANOVA for the linear mixed model of Presence showed that the reliving condition had a significant effect, $F(2, 74) = 21.30, p < .0001$. Pairwise contrasts showed significant differences between ViewVideo and BlurReality ($p = 0.02, t = -2.84$), between ViewVideo and ReliveReality($p < 0.001, t = -6.51$), and between ReliveReality and BlurReality ($p = 0.002, t = -3.63$). These results suggest that participants

Table 1. Exploratory factor analysis (EFA) applied to our questionnaire items, where questions in bold indicate that these items are kept for the final analysis.

No.	Questionnaire items	F1 (PhI)	F2 (PsyI)	F3 (PS)
1	**How satisfied were you when you relived your high-resolution reconstructed experience.**	0.33		**0.38**
2	**In the reconstructed environment I had a sense of "being there."**	0.53		**0.54**
3	**Somehow I felt that the reconstructed environment surrounded me.**	0.31		**0.77**
4	**I felt like I was just perceiving pictures.**			**0.62**
5	**I had a sense of acting in the reconstructed environment, rather than operating something from outside.**			**0.82**
6	How aware were you of the real world surrounding while reliving experience			
7	**I was completely captivated by the reconstructed experience.**			**0.83**
8	**My reconstructed avatar is similar to me**	**0.92**		
9	**I resemble my reconstructed avatar.**	**0.96**		
10	**My reconstructed avatar resembles me.**	**0.87**		
11	**I identify with my reconstructed avatar.**	**0.72**		
12	**My reconstructed avatar is like me in many ways.**	**0.89**		
13	**My reconstructed avatar is an extension of myself.**	**0.95**		
14	**The reconstructed avatar I view reflects who I am.**	**0.71**		
15	**My reconstructed avatar and I are one and the same.**	**0.74**		
16	**The reconstructed avatar I view influences the way I feel about myself.**		**0.57**	
17	**The reconstructed avatar I view is important to my sense of what kind of a person I am.**		**0.67**	
18	**The reconstructed avatar is very special to me.**		**0.84**	
19	**I really care about my reconstructed avatar.**		**0.88**	
20	**I have no emotional connection to my reconstructed avatar**		**0.87**	
21	**I would be heartbroken if I lost my reconstructed avatar.**		**0.86**	
22	**I appreciate my reconstructed avatar.**		**0.68**	
23	**Select the picture which best describes your relationship with your reconstructed avatars.**	**0.68**	0.33	
	SS loadings	6.80	4.58	2.93
	Proportion Variance	0.30	0.20	0.13
	Cumulative Variance	0.30	0.50	0.62

found that the ReliveReality and BlurReality conditions provided higher levels of presence than ViewVideo, and ReliveReality provided the highest levels of presence.

5.3.4 Emotion Ratings. We visually compared the emotions with a radar chart as shown in Figure 6 (d). Overall, participants felt more excited and happy in both ReliveReality and BlurReality. Furthermore, we built linear mixed models for each emotion but only found significant effects of condition on Excited ($F(2, 74) = 3.427, p = .03$) with significant differences between ViewVideo

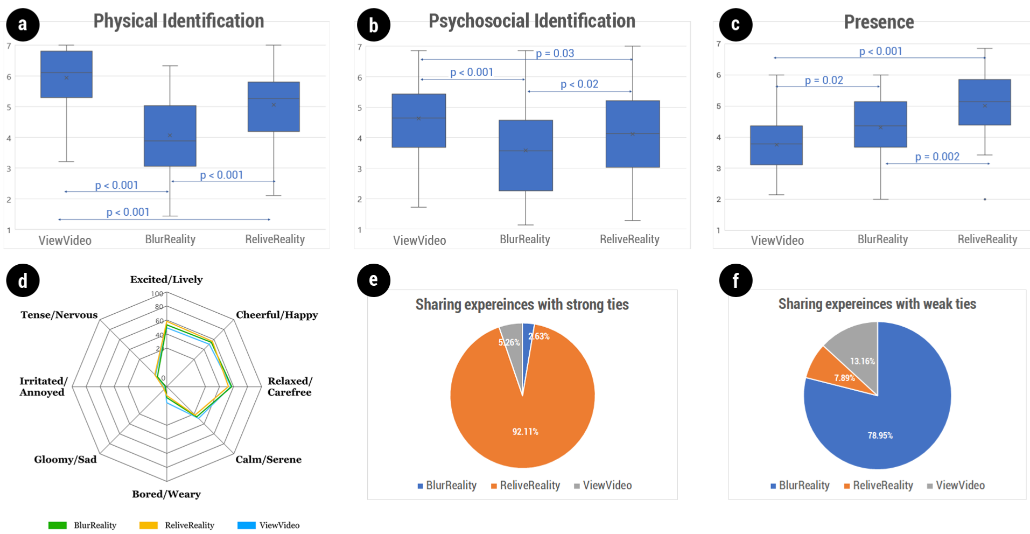

Fig. 6. (a)-(c)boxplots for each factor across reliving conditions and (d) self-reported emotion ratings for each condition. (e) Most preferred reliving method for reliving self experiences (f) Most preferred reliving method for reliving others' experiences

and ReliveReality ($p = 0.03, t = -2.55$). These results suggest that participants felt most excited in ReliveReality, and the blurring effect might decrease the levels of excitement and happiness.

5.3.5 Preference. As shown in Figure 6 (e) and (f), when asked about the most preferred method for sharing experiences with people that close to them, 35 out of 38 participants (92%) chose ReliveReality, 2 (5%) participants chose ViewVideo and only 1 participant (3%) chose BlurReality. In terms of the preferred sharing method for weak ties, 30 out of 38 (79%) participants selected BlurReality, 5 (13%) participants selected ViewVideo and 3 participants (8%) selected ReliveReality. These results indicate that participants prefer to share highly detailed reconstructed experiences with people who they know well, but with their weak ties, they prefer sharing obfuscated, blurred versions to reduce privacy concerns.

5.4 Sharing Behavior Analysis

We first looked into overall comfort with sharing each type of content across all recipients. We included the average "comfort with sharing" as the dependent variable and added fixed effects of condition and previous social media usage. We included participant ID as a random effect. There were significant main effects of type of reliving contents ($F(2, 74) = .002$), along with previous social media usage, which was marginally significant ($F(1, 36) = .054$), reflecting the fact that content type affected comfort with sharing differently. Furthermore, pairwise comparison showed significant differences between ViewVideo and BlurReality ($p = 0.002, t = -3.51$) and between ReliveReality and BlurReality ($p = 0.04, t = 0.307$), but not between ViewVideo and ReliveReality ($p = 0.57$). These results indicate that people are more comfortable sharing reconstructed experiences with blurring effects than recorded videos and original reconstructed experiences.

Figure 7 demonstrates participants' comfort sharing with each recipient for all three different reliving contents in ViewVideo, ReliveReality and BlurReality. We see different patterns for people who are close to the sharer compared to people who are not close. Thus, to better understand

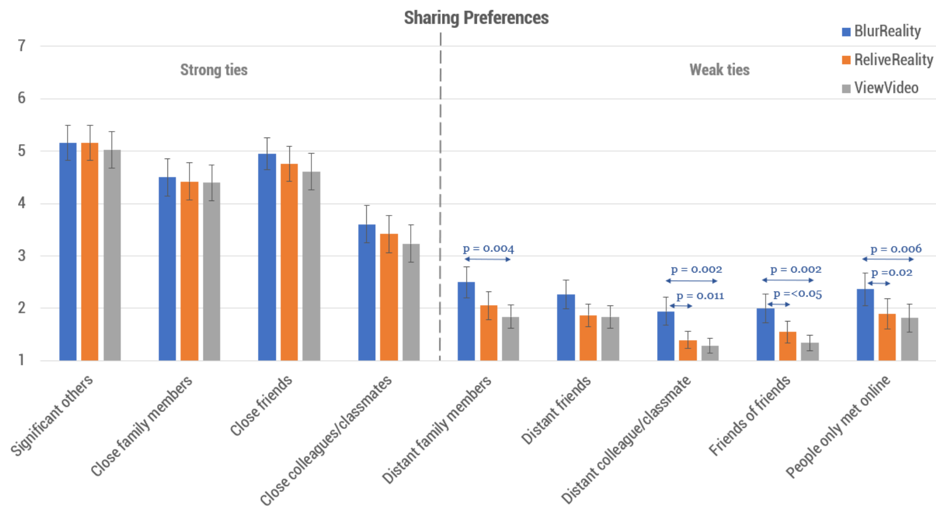

Fig. 7. Participants' likelihood to share with each recipient for all types of reliving contents including recorded videos, reconstructed experiences and reconstructed experiences with blurring obfustication.

how different types of reliving content influence comfort with sharing, we built linear mixed models for each recipient. Similar to the model above, we included comfort with sharing as the dependent variable and added fixed effects of content type and previous social media usage. As shown in the Table 2, we found significant effects of content type on comfort with sharing across multiple types of recipients including distant family members ($F(2, 74) = 5.66, p = .005$), distant colleague/classmates ($F(2, 74) = 7.24, p = .001$), friends of friends ($F(2, 74) = 6.51, p = .002$) and online friends ($F(2, 74) = 5.90, p = .004$) , as well as marginal significance in distant friends ($F(2, 74) = 2.98, p = .056$). These results suggest that people are comfortable sharing all types of reliving contents to people who they know well, but people are less likely to share recorded video or 3D reconstructed experiences to recipients with whom they are not close. However, applying the blurring effect on reconstructed experience significantly increased the comfortability of sharing when recipients are not close.

5.5 Qualitative Analysis Approach

We analyzed participant responses using an iterative approach inspired by grounded theory [73] and thematic analysis [6]. Using the transcribed audio segments of interest, two researchers separately developed a set of codes that covered various aspects of sharing reconstructed experiences such as quality of sharing, privacy concerns for reconstructed avatars and environments, and sharing behaviors for strong ties and weak ties. Using the codes as analytic lenses, we clustered various aspects of the experience under emerging themes, which shed light on how differences between the three sharing conditions. Until both researchers reached full consensus, we iteratively distilled the most relevant themes into insights that we present below in the discussion alongside the quantitative data.

Table 2. ANOVA results on each recipent's linear mixed model and pairwise comparison with Tukey adjustments results

Recipients	ANOVA results	Pairwise comparison
Significant Others	$F_{2,102} = 31.30, p < .001$	No significance
Close family members	$F_{2,102} = 31.30, p < .001$	No significance
Close friends	$F_{2,102} = 31.30, p < .001$	No significance
Close colleague/classmate	$F_{2,102} = 31.30, p < .001$	No significance
Distant family members	$F_{2,74} = 5.66, p = .005$	BlurReality > ViewVideo ($p = .004$)
Distant friends	$F_{2,74} = 2.98, p = .056$	Approached significance
Distant colleague/classmate	$F_{2,74} = 7.24, p = .001$	BlurReality > ViewVideo ($p = .002$) BlurReality > ReliveReality ($p = .011$)
Friends of friends	$F_{2,74} = 6.51, p = .002$	BlurReality > ViewVideo ($p = .002$) BlurReality > ReliveReality ($p < .05$)
People only met online	$F_{2,74} = 5.90, p = .004$	BlurReality > ViewVideo ($p = .006$) BlurReality > ReliveReality ($p = .02$)

6 DISCUSSION

6.1 ReliveReality is preferred for strong ties while BlurReality mitigates privacy concerns for weak ties

Regarding *sharing* their own reconstructed content, participants felt comfortable sharing all three types with people that they are close to. However, we note that if participants were asked which condition they would prefer to *use* when sharing to their strong ties, they overwhelmingly (92%) selected the highly detailed ReliveReality experience as shown in Figure 6 (e) because of higher levels of presence and the ability to move around in the environment.

As P29 participant puts it:

> If I'm going to share it with my family or close friends, all three [conditions] would be fine to share. But I prefer to share the 3D clear one [ReliveReality], because that allows them to know more details about the experience. They can move around in my environment and not just look at the video from a fixed point.

However, as shown in Figure 7, participants felt more comfortable to sharing blurred reconstructed experiences with recipients that they don't know well (e.g. distant friends, distant classmates, online friends) compared to either the video or the highly detailed reconstructed experience. In addition, 80% of participants rated BlurReality as their preferred sharing method when sharing experiences to weak ties as shown in Figure 6 (f) .

Participants indicated that they identified more with the high physical similarity avatars in ReliveReality and this identification made them less likely to share it with recipients who are not close to them due to increased privacy concerns. On the other hand, the blurred avatars and environments in the BlurReality reduced the chance of being recognized or disclosing sensitive content in the environment. As a result, participants were more likely to share the BlurReality with their weak ties over the other two conditions.

As P38 participant puts it:

> If I'm going to share it with my family or close friends, all three [conditions] would be fine to share. But for the people I don't know very well, I will only share the blurred version [BlurReality] since I don't really want to let other people see me dancing since I'm not a good dancer. So if it's blurry, then I think it would be fine to do it since they won't know who the person is.

Even though 79% of participants rated BlurReality as their preferred sharing methods for weak ties, participants overall showed limited desires to share any of the three types with recipients that they don't know well. However, prior studies have shown that users are unlikely to have accurate understanding and control of their audience when sharing content on Social Network Sites (SNSs). In the case where users might inadvertently share content, our study results suggested that obfuscating reconstructed experience can be an effective strategy for enhancing privacy when shared reconstructed experiences are viewed by undesired audiences.

6.2 Reconstructed avatars reduce identification but maintain perceived privacy

The realism results indicated that the ReliveReality prototype can generate avatars resembling participants and reconstruct the 3D environments in great detail. However, compared with the ViewVideo condition, participants reported lower physical and psychosocial identification with their reconstructed avatars than with their video representations. The qualitative data from the interview helps us identify two main factors that reduced avatar identification.

First of all, 80% of all participants indicated that the "reconstruction process" itself and being in a virtual environment made the reconstructed avatar seem less real, resulting in less avatar identification.

As P10 participant puts it:

> When I viewed the video, I knew it's me for sure because I recorded the video and I still remember what I did. But the avatar was created by someone or some programs and it was being shown in a virtual world. Even though the avatar really looks like me. I can even see my mole on its face. But I still view it as a virtual character, not the real me.

Another reason is that 30% of the participants mentioned that the avatar's dance movements are not as smooth as the movements in the video, especially for rapid movements in different dances.

As P2 participant puts it:

> This is the coolest avatar [ReliveReality] I have ever seen. It resembles me in almost every detail like the hair, logo on the clothes, even my slippers. The only thing that made me feel less real was its movement. Most of the time it looked great and natural but sometimes it became jittery and less real especially around the feet.

Despite the fact that the avatar identification was lower in the ReliveReality condition, participants indicated that they still related themselves to their reconstructed avatars and viewed the reconstructed avatars as part of themselves. For example, to investigate how participants identified with their reconstructed avatars, we asked whether they would feel uncomfortable if someone approached very close and poked at their reconstructed avatars during the interview. 92% of participants reported that they would feel uncomfortable if their reconstructed avatars were being treated in a manner that would be inappropriate in the physical world.

As participant P31 puts it:

> I know the avatar is a virtual character, but it's a recreation of me. It looks like me, dresses like me and moves like me. So I still feel connected with it. If someone disrespects my avatar, I will also feel offended by that person.

P22 participant also mentions that:

> It depends on who the person is. If he or she is someone really close to me like my best friend, I probably will feel ok since I might want to have fun with their avatars too. But if it's someone I don't know well especially if he is male, I will feel really uncomfortable.

Furthermore, identifying with reconstructed avatars also made participants less likely to share reconstructed experiences with someone they don't know.

As P9 participant puts it:

> I feel embarrassed if other people [someone I don't know well] see me performing funny dances. So I probably won't share the video with them. For the avatar [ReliveReality], I won't share it with them either because the avatar looks like me and it's like part of me.

6.3 Privacy concerns about reconstructed environments

In addition to reconstructed avatars, participants also expressed privacy concerns about their reconstructed environments, leading them less likely to share it with socially distant recipients as shown in Figure 7. Since the reconstructed environments in ReliveReality allow recipients to move closer and view the details of the environment, participants reported that it would raise more privacy concerns.

As P9 participant puts it:

> Maybe I film a video of my room and then people can just interact with everything in my room so I felt that might be a little bit of an invasion of privacy. I feel uncomfortable that they can see some...I don't know… books in my bookshelves or the posters of my favorite idols or whatever.

However, we didn't find a significant difference in sharing preferences between ViewVideo and ReliveReality to socially distant recipients. As pointed out by several participants in the interview, one plausible reason is that they were aware that the environments will be 3D reconstructed, so they had made sure that there were no sensitive contents in the environment before they recorded the video. This can be viewed as avoidance behavior described in the behavioral privacy model [14] and it illustrates the importance of notifying users what contents will be captured and reconstructed beforehand to reduce privacy concerns.

As P37 participant put it:

> I want to make it [my room] neat and tidy before I record it. So I cleaned up my room and hid some personal stuff before the study. I didn't want to show some of my personal stuff.

Some participants also suggested that reducing privacy concerns would be more critical in social spaces compared to the personal space (i.e. their rooms) in the study.

As P11 participant put it:

> If I go to a party at my friend's house and I record it, there might be some stuff that my friend doesn't want to show to others. But I won't be able to control it since it's not my room. I think blurring would be even more useful in this case.

P23 participant also mentioned that:

> I can imagine the blurring will be really helpful if I want to reconstruct a social space like a coffee shop or park. There will be some people in the background and I think this technique [ReliveReality] will also reconstruct those people and their personal stuff. That might violate their privacy. So I will want to blur those [reconstructed] people and stuff before I share it [reconstructed experience].

Our study results reveal that participants had privacy concerns when sharing reconstructed environments and different physical environments would change users' perceived privacy depending on the potential sensitive content in the environments. In addition, the study results also suggested that obfuscation techniques like blurring can mitigate these privacy concerns.

6.4 Blurring preserves privacy but reduces presence and thus satisfication

In the BlurReality condition, both facial and clothing details of the reconstructed avatars were obfuscated by blurring. This made avatars look less like the participants, resulting in significantly lower physical and psychosocial identification than avatars in ReliveReality condition. Blurring also preserves privacy by preventing recognition, so participants indicated that they were more willing to share BlurReality contents to people that are not close.

As P13 participant puts it:

> I kind of like the blurred version [BlurReality] even more than the clear one [ReliveReality]. I feel a bit embarrassed by looking at myself dance awkwardly and I definitely don't want to let other ppl see it. But the blurred avatar hides my identity and it also make me feel like an oil painting style. So I think I will be comfortable sharing this to people I don't know well.

However, even though participants can still move around in the blurred reconstructed environment, blurring the environment significantly decreased the levels of presence, resulting in a less satisfactory reliving experience. Therefore, futher research is needed to explore other obfuscation techniques that can enhance privacy without sacrificing too much of the viewing experience.

6.5 Limitations and applicability of our findings

There are multiple limitations to our study. All participants in our study were undergraduate students. This population may be more tech-savvy as well as more knowledgeable about privacy issues than some other populations. Some participants (38%) mentioned that ReliveReality is cool and new, implying that ReliveReality might produce novelty effects in some participants. In addition, our study examined a particular context in which users mimicked funny dance movements and shared those reconstructed experience with others. We know from previous research [45] that sharing different kinds of content can raise different privacy concerns.

In order to mitigate these issues, we applied a mixed-method approach and conducted semi-structured interviews to identify what aspects of ReliveReality made participants prefer it, and how reconstructed experiences raised privacy concerns. For example, our interview results suggested that the ability to move around in the reconstructed environments and relating themselves with their reconstructed avatars make them prefer to share experiences via ReliveReality. We also gained better understanding of their privacy concerns on ReliveReality (e.g. identifying with avatars, sensitive content in reconstructed environments), leading them to prefer BlurReality when sharing content to weak-tie recipients. Understanding these underlying reasons helps us mitigate the novelty effect and consider how these findings might generalize to other populations. Although further study is needed to increase the external validity, our findings and implications can guide future research to investigate the sharing behaviors and privacy concerns of reconstructed experience for other populations and shared content.

7 DESIGN IMPLICATIONS

As suggested by the behavioral privacy model, we found that participants' sharing preferences and privacy concerns were affected by both the types of reliving contents (i.e. ViewVideo, ReliveReality and BlurReality) and the recipients. The idea of *sharing* this reconstructed content raised critical privacy concerns and created challenges which could affect the quality of the sharing experience. We synthesize the ideas that arise from our study, and organize design implications by two core elements in the behavioral privacy model – controlling content and controlling recipients.

7.1 Controlling recipients - Enabling selective sharing for 3D reconstructed contents

When sharing photos or videos, the current recipient controlling tools offered by SNSs allow users to to disclose their personal information selectively. This is a common avoidance behavior described in the behavioral theory of privacy that people use to avoid privacy leakage before it occurs. Based on our results, when sharing personal 3D reconstructed contents with a high level of detail, participants have similar but increased privacy concerns compared to sharing videos. However, most of the 3D model sharing platforms such as Sketchfab, p3d.in or Display.land app only allow users to choose between private or public options (i.e. anyone with the link can access it). Making content private will defeat some of the motivations for sharing, such as providing entertaining content to others, strengthening social connections or simply for the self-fulfillment of sharing [74]. However, making content public will increase the privacy risks. Therefore, the system should provide better fine-grained recipient control mechanisms to allow people to selectively share their 3D reconstructed contents.

7.2 Controlling content- Adapting reconstructed contents based on recipients

As suggested by the behavioral theory of privacy, different recipients will influence the perceived privacy even when sharing the same experience. According to our study results, participants prefer to share highly detailed reconstructed experiences with their strong ties, but share blurred reconstructed experiences with their weak ties. Thus, instead of sharing either highly detailed or blurred reconstructed experiences to all recipient groups, it's critical to design techniques that adapt 3D reconstructed content based on different recipients. For example, closer friends can view more details than strangers when they view the same reconstructed experience.

7.3 Controlling content- Preserving privacy while maintaining satisfaction

With the focus on the content element in the behavioral theory of privacy, obscuring part of the content to protect privacy has been widely adopted by many existing photo or video applications. However, none of these mechanisms have been introduced in the context of sharing 3D reconstructed content. In our study, we used one of the most common obfuscation methods, blurring, for both the reconstructed avatars and whole environments. Our study results suggest that blurring can significantly decrease avatar identification and reduce privacy concerns, but it also significantly reduces presence levels, resulting in a less satisfactory reliving experience. To address these issues, we should carefully design techniques to identify the sensitive content or private regions in the environment and only apply obfuscation to those regions instead of blurring the whole 3D environment. Besides, in addition to blurring, we should also explore more obfuscation methods that would be particularly suitable to 3D content such as morphing the self- avatar to a generic avatar or replacing the texture of a sensitive 3D object (e.g. a 3D photo frame showing the photos with family) to a generic texture (e.g. a generic nature photo) that could potentially obfuscate sensitive contents without degrading the reconstructed experience.

7.4 Customizable 3D reconstructed contents

The concept of transformed social interaction [8] has demonstrated that strategic decoupling of rendered forms and behaviors from the actual ones in the physical world allows interactants to break many constraints that are inherent in face-to-face interaction. When physical experiences are recorded and reconstructed into virtual worlds, we can transform the representations of the reconstructed avatar and environments to better fit to various scenarios rather than rendering exactly the same representations as the physical world.

For example, users sometimes create their avatars in the likeness of their ideal self more than their actual self [10, 21]. They might also emphasize certain characteristics and conceal others in their avatars in order to manage their impressions on others [26]. Several participants reported that they'd like to customize their reconstructed avatars such as making the avatar taller, making the skin brighter or changing the clothes of their reconstructed avatars during the interview. In addition to wishing to customize their reconstructed avatars, participants also wished to alter the reconstructed environment, for example by removing some objects in the scene or rearranging the furniture layout. Therefore, we should leverage the unique possiblites of the virtual world and allow users to customize their own reconstructed experiences.

8 FUTURE WORK

In addition to the limitations arising from population and study design discussed above, the current prototype and study has some technical limitations that should be discussed. Although the framework of ReliveReality only relies on a RGB camera, the types of experiences that can be reconstructed are limited. The 3D human pose estimation only allows us to relive user's motions but not any moving objects in the environment. In addition, 3D environment reconstruction from videos is time-consuming and will generate unreliable results in some common situations such as textureless walls and reflective floors and tables. In the future, we plan to integrate other techniques based on RGB-D cameras to address these limitations and increase the quality of reconstructed experiences when users have the required hardware.

As discussed above, our study results might be affected by the study population and the particular shared content (i.e. funny dance movements). Further research in different contexts is needed to validate the generalizability of our findings. The current ReliveReality prototype already supports reliving reconstructed experiences in virtual reality (VR). In the future, we plan to investigate how higher presence and sense of embodiment in more immersive VR would affect our study results. Lastly, we also plan to explore how different obfuscation methods affect the reconstructed experience and implement different techniques to allow people better customize their reconstructed experience. In addition to the technical improvements, we'd like to also explore how different modalities such as audios and haptics can be reconstructed and utilized for a better reliving experience. In terms of sharing, we're also interested in investigating how to preserve privacy when there are multiple people in the same reconstructed experience such as a reconstructed party experience or a reconstructed basketball game.

9 CONCLUSION

In this paper, we present ReliveReality, a new experience-sharing method that allows people to reconstruct their experiences in 3D and share them easily with others. ReliveReality utilized various deep-learning based computer vision techniques to reconstruct users, their actions, and their environment using only a single RGB camera. Through ReliveReality, people can enter into other's reconstructed experiences and view the experience from any perspective. To understand how people relate to their reconstructed avatars, and how such new reconstructed experiences influence people's sharing preference and privacy concerns, we grounded a user study in the behavioral theory of privacy and conducted a 1X3 two-stage within-subject study with 44 participants. We first utilized ReliveReality to generate reconstructed experiences for all participants including participants' avatars, movements, and their environments. Then participants relived their experiences through three different reliving contents including: 1) recorded videos, 2) reconstructed experiences, and 3) reconstructed experience with blurring obfuscation. Our results suggested that participants identified less with their reconstructed avatars compared with their video representations, but they still relate with their avatars. Because of the ability to relive the experiences from different perspectives, they prefer to share highly detailed reconstructed experiences with with strong tie recipients such as close friends and family members, but less likely to share them with people they don't know well. Although applying the blurring obfuscation on reconstructed experiences mitigate the privacy concerns and significantly increased the likelihood of sharing them with the weak-tie recipients, it also degraded the viewer's levels of presence and satisfaction significantly. Based on our study result and informed by the temporal aspects of privacy in the behavioral theory of privacy, we propose design implications to better preserve privacy while sharing reconstructed experiences, since these are experiences that will arise more often as technology advances.

REFERENCES

[1] 2018. AliceVision: Photogrammetry Pipeline. https://alicevision.org/photogrammetry.

[2] Ilya Afanasyev, Artur Sagitov, and Evgeni Magid. 2015. ROS-based SLAM for a Gazebo-simulated mobile robot in image-based 3D model of indoor environment. In *International Conference on Advanced Concepts for Intelligent Vision Systems*. Springer, 273–283.

[3] Sun Joo-Grace Ahn and Jeremy Bailenson. 2011. Self-endorsing versus other-endorsing in virtual environments: the effect on brand attitude and purchase intention. *Journal of Advertising* 40 (01 2011), 93–106. https://doi.org/10.2307/23048707

[4] Thiemo Alldieck, Marcus Magnor, Bharat Lal Bhatnagar, Christian Theobalt, and Gerard Pons-Moll. 2019. Learning to reconstruct people in clothing from a single RGB camera. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 1175–1186.

[5] Thiemo Alldieck, Marcus Magnor, Weipeng Xu, Christian Theobalt, and Gerard Pons-Moll. 2018. Detailed human avatars from monocular video. In *2018 International Conference on 3D Vision (3DV)*. IEEE, 98–109.

[6] J. Aronson. 1994. A pragmatic view of thematic analysis. *The Qualitative Report* 2 (1994).

[7] Laura Aymerich-Franch, René Kizilcec, and Jeremy Bailenson. 2014. The Relationship between Virtual Self Similarity and Social Anxiety. *Frontiers in human neuroscience* 8 (11 2014), 944. https://doi.org/10.3389/fnhum.2014.00944

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

[9] Jaime Banks and Nicholas Bowman. 2016. Emotion, anthropomorphism, realism, control: Validation of a merged metric for player–avatar interaction (PAX). *Computers in Human Behavior* 54 (01 2016), 215–223. https://doi.org/10.1016/j.chb.2015.07.030

[10] Katherine Bessière, A Seay, and Sara Kiesler. 2007. The Ideal Elf: Identity Exploration in World of Warcraft. *Cyberpsychology & behavior : the impact of the Internet, multimedia and virtual reality on behavior and society* 10 (09 2007), 530–5. https://doi.org/10.1089/cpb.2007.9994

[11] Federica Bogo, Michael J Black, Matthew Loper, and Javier Romero. 2015. Detailed full-body reconstructions of moving people from monocular RGB-D sequences. In *Proceedings of the IEEE International Conference on Computer Vision*. 2300–2308.

[12] Federica Bogo, Angjoo Kanazawa, Christoph Lassner, Peter Gehler, Javier Romero, and Michael J Black. 2016. Keep it SMPL: Automatic estimation of 3D human pose and shape from a single image. In *European Conference on Computer Vision*. Springer, 561–578.

[13] Michael Boyle, Christopher Edwards, and Saul Greenberg. 2000. The Effects of Filtered Video on Awareness and Privacy. In *Proceedings of the 2000 ACM Conference on Computer Supported Cooperative Work* (Philadelphia, Pennsylvania, USA) *(CSCW '00)*. Association for Computing Machinery, New York, NY, USA, 1–10. https://doi.org/10.1145/358916.358935

[14] Kelly E. Caine. 2009. Exploring everyday privacy behaviors and misclosures. In *Ph.D. Dissertation*. Georgia Institute of Technology.

[15] Zhe Cao, Gines Hidalgo, Tomas Simon, Shih-En Wei, and Yaser Sheikh. 2019. OpenPose: realtime multi-person 2D pose estimation using Part Affinity Fields. *IEEE transactions on pattern analysis and machine intelligence* 43, 1 (2019), 172–186.

[16] Zhe Cao, Tomas Simon, Shih-En Wei, and Yaser Sheikh. 2017. Realtime multi-person 2d pose estimation using part affinity fields. In *Proceedings of the IEEE conference on computer vision and pattern recognition*. 7291–7299.

[17] Ching-Hang Chen and Deva Ramanan. 2017. 3d human pose estimation= 2d pose estimation+ matching. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 7035–7043.

[18] Jonathan Cohen. 2001. Defining identification: A theoretical look at the identification of audiences with media characters. *Mass communication & society* 4, 3 (2001), 245–264.

[19] AB Costello and Osborne Jason. 2005. Best Practices in Exploratory Factor Analysis: Four Recommendations for Getting the Most From Your Analysis. *Practical Assessment, Research & Evaluation* 10 (01 2005), 1–9.

[20] Rishabh Dabral, Anurag Mundhada, Uday Kusupati, Safeer Afaque, Abhishek Sharma, and Arjun Jain. 2018. Learning 3d human pose from structure and motion. In *Proceedings of the European Conference on Computer Vision (ECCV)*. 668–683.

[21] Nicolas Ducheneaut, Ming-Hui Wen, Nicholas Yee, and Greg Wadley. 2009. Body and Mind: A Study of Avatar Personalization in Three Virtual Worlds *(CHI '09)*. Association for Computing Machinery, New York, NY, USA, 1151–1160. https://doi.org/10.1145/1518701.1518877

[22] Hongchao Fan, Wei Yao, and Qing Fu. 2014. Segmentation of sloped roofs from airborne LiDAR point clouds using ridge-based hierarchical decomposition. *Remote Sensing* 6, 4 (2014), 3284–3301.

[23] Hao-Shu Fang, Shuqin Xie, Yu-Wing Tai, and Cewu Lu. 2017. Rmpe: Regional multi-person pose estimation. In *Proceedings of the IEEE International Conference on Computer Vision*. 2334–2343.

[24] Jesse Fox and Jeremy Bailenson. 2009. Virtual Self-Modeling: The Effects of Vicarious Reinforcement and Identification on Exercise Behaviors. *Media Psychology - MEDIA PSYCHOL* 12 (03 2009), 1–25. https://doi.org/10.1080/15213260802669474

[25] Guo Freeman and Divine Maloney. 2021. Body, Avatar, and Me: The Presentation and Perception of Self in Social Virtual Reality. *Proceedings of the ACM on Human-Computer Interaction* 4, CSCW3 (2021), 1–27.

[26] Erving Goffman. 1959. *The Presentation of Self in Everyday Life*. Anchor.

[27] Mar Gonzalez-Franco, Zelia Egan, Matthew Peachey, Angus Antley, Tanmay Randhavane, Payod Panda, Yaying Zhang, Cheng Yao Wang, Derek F Reilly, Tabitha C Peck, et al. 2020. MoveBox: Democratizing MoCap for the Microsoft Rocketbox Avatar Library. In *2020 IEEE International Conference on Artificial Intelligence and Virtual Reality (AIVR)*. IEEE, 91–98.

[28] Mar Gonzalez-Franco, Anthony Steed, Steve Hoogendyk, and Eyal Ofek. 2020. Using facial animation to increase the enfacement illusion and avatar self-identification. *IEEE transactions on visualization and computer graphics* 26, 5 (2020), 2023–2029.

[29] M.S. Granovetter. 1973. The Strength of Weak Ties. 78, 6 (1973), 1360–1380.

[30] Pranay Gupta, Anirudh Thatipelli, Aditya Aggarwal, Shubh Maheshwari, Neel Trivedi, Sourav Das, and Ravi Kiran Sarvadevabhatla. 2020. Quo Vadis, Skeleton Action Recognition ? arXiv:2007.02072 [cs.CV]

[31] Rakibul Hasan, Yifang Li, Eman Hassan, Kelly Caine, David J. Crandall, Roberto Hoyle, and Apu Kapadia. 2019. Can Privacy Be Satisfying? On Improving Viewer Satisfaction for Privacy-Enhanced Photos Using Aesthetic Transforms. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3290605.3300597

[32] Carrie Heeter. 1992. Being there: The subjective experience of presence. *Presence: Teleoperators & Virtual Environments* 1, 2 (1992), 262–271.

[33] Zeng Huang, Yuanlu Xu, Christoph Lassner, Hao Li, and Tony Tung. 2020. ARCH: Animatable Reconstruction of Clothed Humans. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 3093–3102.

[34] Scott E. Hudson and Ian Smith. 1996. Techniques for Addressing Fundamental Privacy and Disruption Tradeoffs in Awareness Support Systems. In *Proceedings of the 1996 ACM Conference on Computer Supported Cooperative Work* (Boston, Massachusetts, USA) *(CSCW '96)*. Association for Computing Machinery, New York, NY, USA, 248–257. https://doi.org/10.1145/240080.240295

[35] Catalin Ionescu, Dragos Papava, Vlad Olaru, and Cristian Sminchisescu. 2013. Human3. 6m: Large scale datasets and predictive methods for 3d human sensing in natural environments. *IEEE transactions on pattern analysis and machine intelligence* 36, 7 (2013), 1325–1339.

[36] Sasa Junuzovic, Kori Inkpen, John Tang, Mara Sedlins, and Kristie Fisher. 2012. To see or not to see: a study comparing four-way avatar, video, and audio conferencing for work. In *Proceedings of the 17th ACM international conference on Supporting group work*. 31–34.

[37] Sanjay Kairam, Mike Brzozowski, David Huffaker, and Ed Chi. 2012. Talking in Circles: Selective Sharing in Google+. Association for Computing Machinery, New York, NY, USA.

[38] Sanjay Kairam, Joseph 'Jofish' Kaye, John Alexis Guerra-Gomez, and David A. Shamma. 2016. Snap Decisions? How Users, Content, and Aesthetics Interact to Shape Photo Sharing Behaviors *(CHI '16)*. Association for Computing Machinery, New York, NY, USA, 113–124. https://doi.org/10.1145/2858036.2858451

[39] Angjoo Kanazawa, Jason Y Zhang, Panna Felsen, and Jitendra Malik. 2019. Learning 3d human dynamics from video. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 5614–5623.

[40] Zhizhong Kang, Juntao Yang, Zhou Yang, and Sai Cheng. 2020. A Review of Techniques for 3D Reconstruction of Indoor Environments. *ISPRS International Journal of Geo-Information* 9, 5 (2020), 330.

[41] Sung Yeun Kim. 2014. "That's me in the game": Being present and identifying avatar. In *Ph.D. Dissertation*. Syracuse University.

[42] Muhammed Kocabas, Nikos Athanasiou, and Michael J Black. 2020. VIBE: Video inference for human body pose and shape estimation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 5253–5263.

[43] Jordan Koulouris, Zoe Jeffery, James Best, Eamonn O'Neill, and Christof Lutteroth. 2020. Me vs. Super (wo) man: Effects of Customization and Identification in a VR Exergame. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–17.

[44] Dong Dong Li, Albert Kien Liau, and Angeline Khoo. 2013. Player–Avatar Identification in video gaming: Concept and measurement. *Computers in Human Behavior* 29, 1 (2013), 257–263.

[45] Yifang Li, Nishant Vishwamitra, Hongxin Hu, and Kelly Caine. 2020. Towards A Taxonomy of Content Sensitivity and Sharing Preferences for Photos *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3313831.3376498

[46] Yifang Li, Nishant Vishwamitra, Bart P. Knijnenburg, Hongxin Hu, and Kelly Caine. 2017. Effectiveness and Users' Experience of Obfuscation as a Privacy-Enhancing Technology for Sharing Photos. *Proc. ACM Hum.-Comput. Interact.* 1, CSCW, Article 67 (Dec. 2017), 24 pages. https://doi.org/10.1145/3134702

[47] Yabing Liu, Krishna P. Gummadi, Balachander Krishnamurthy, and Alan Mislove. 2011. Analyzing Facebook privacy settings: User expectations vs. reality. *Proceedings of the ACM SIGCOMM Internet Measurement Conference, IMC* (11 2011). https://doi.org/10.1145/2068816.2068823

[48] Matthew Loper, Naureen Mahmood, Javier Romero, Gerard Pons-Moll, and Michael J Black. 2015. SMPL: A skinned multi-person linear model. *ACM transactions on graphics (TOG)* 34, 6 (2015), 1–16.

[49] Julieta Martinez, Rayat Hossain, Javier Romero, and James J Little. 2017. A simple yet effective baseline for 3d human pose estimation. In *Proceedings of the IEEE International Conference on Computer Vision*. 2640–2649.

[50] Dushyant Mehta, Oleksandr Sotnychenko, Franziska Mueller, Weipeng Xu, Mohamed Elgharib, Pascal Fua, Hans-Peter Seidel, Helge Rhodin, Gerard Pons-Moll, and Christian Theobalt. 2020. XNect: Real-time multi-person 3D motion capture with a single RGB camera. *ACM Transactions on Graphics (TOG)* 39, 4 (2020), 82–1.

[51] Dushyant Mehta, Srinath Sridhar, Oleksandr Sotnychenko, Helge Rhodin, Mohammad Shafiei, Hans-Peter Seidel, Weipeng Xu, Dan Casas, and Christian Theobalt. 2017. Vnect: Real-time 3d human pose estimation with a single rgb camera. *ACM Transactions on Graphics (TOG)* 36, 4 (2017), 1–14.

[52] Edward M Mikhail, James S Bethel, and J Chris McGlone. 2001. Introduction to modern photogrammetry. *New York* (2001), 19.

[53] Emily Moyer-Gusé, Adrienne H Chung, and Parul Jain. 2011. Identification with characters and discussion of taboo topics after exposure to an entertainment narrative about sexual health. *Journal of communication* 61, 3 (2011), 387–406.

[54] Richard A Newcombe, Dieter Fox, and Steven M Seitz. 2015. Dynamicfusion: Reconstruction and tracking of non-rigid scenes in real-time. In *Proceedings of the IEEE conference on computer vision and pattern recognition*. 343–352.

[55] Richard A Newcombe, Shahram Izadi, Otmar Hilliges, David Molyneaux, David Kim, Andrew J Davison, Pushmeet Kohi, Jamie Shotton, Steve Hodges, and Andrew Fitzgibbon. 2011. KinectFusion: Real-time dense surface mapping and tracking. In *2011 10th IEEE International Symposium on Mixed and Augmented Reality*. IEEE, 127–136.

[56] Helen Nissenbaum. 2004. Privacy As Contextual Integrity. *Washington Law Review* 79 (05 2004).

[57] Anne Oeldorf-Hirsch and S. Shyam Sundar. 2016. Social and Technological Motivations for Online Photo Sharing. *Journal of Broadcasting and Electronic Media* 60 (10 2016), 624–642. https://doi.org/10.1080/08838151.2016.1234478

[58] D. Parmar, J. Isaac, S. V. Babu, N. D'Souza, A. E. Leonard, S. Jörg, K. Gundersen, and S. B. Daily. 2016. Programming moves: Design and evaluation of applying embodied interaction in virtual environments to enhance computational thinking in middle school students. In *2016 IEEE Virtual Reality (VR)*. 131–140.

[59] Dario Pavllo, Christoph Feichtenhofer, David Grangier, and Michael Auli. 2019. 3d human pose estimation in video with temporal convolutions and semi-supervised training. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 7753–7762.

[60] Sandra Petronio. 2015. *Communication Privacy Management Theory*. American Cancer Society, 1–9. https://doi.org/10.1002/9781118540190.wbeic132 arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1002/9781118540190.wbeic132

[61] Gerard Pons-Moll, Sergi Pujades, Sonny Hu, and Michael J Black. 2017. ClothCap: Seamless 4D clothing capture and retargeting. *ACM Transactions on Graphics (TOG)* 36, 4 (2017), 1–15.

[62] Yasmeen Rashidi, Tousif Ahmed, Felicia Patel, Emily Fath, Apu Kapadia, Christena Nippert-Eng, and Norman Makoto Su. 2018. "You don't want to be the next meme": College Students' Workarounds to Manage Privacy in the Era of Pervasive Photography. In *Fourteenth Symposium on Usable Privacy and Security (SOUPS 2018)*. USENIX Association, Baltimore, MD, 143–157. https://www.usenix.org/conference/soups2018/presentation/rashidi

[63] Mir Rayat Imtiaz Hossain and James J Little. 2018. Exploiting temporal information for 3d human pose estimation. In *Proceedings of the European Conference on Computer Vision (ECCV)*. 68–84.

[64] H. Regenbrecht and T. Schubert. 2002. Real and Illusory Interactions Enhance Presence in Virtual Environments. *Presence* 11, 4 (2002), 425–434.
[65] Nadia Robertini, Dan Casas, Edilson De Aguiar, and Christian Theobalt. 2017. Multi-view performance capture of surface details. *International Journal of Computer Vision* 124, 1 (2017), 96–113.
[66] Wallace Sadowski and Kay Stanney. 2002. Presence in virtual environments. (2002).
[67] Shunsuke Saito, Zeng Huang, Ryota Natsume, Shigeo Morishima, Angjoo Kanazawa, and Hao Li. 2019. Pifu: Pixel-aligned implicit function for high-resolution clothed human digitization. In *Proceedings of the IEEE International Conference on Computer Vision*. 2304–2314.
[68] Shunsuke Saito, Tomas Simon, Jason Saragih, and Hanbyul Joo. 2020. PIFuHD: Multi-Level Pixel-Aligned Implicit Function for High-Resolution 3D Human Digitization. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 84–93.
[69] Thomas Schubert, Frank Friedmann, and Holger Regenbrecht. 2001. The Experience of Presence: Factor Analytic Insights. *Presence* 10 (06 2001), 266–281. https://doi.org/10.1162/105474601300343603
[70] Valentin Schwind, Pascal Knierim, Nico Haas, and Niels Henze. 2019. Using Presence Questionnaires in Virtual Reality *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3290605.3300590
[71] Ari Shapiro, Andrew Feng, Ruizhe Wang, Hao Li, Mark Bolas, Gerard Medioni, and Evan Suma. 2014. Rapid avatar capture and simulation using commodity depth sensors. *Computer Animation and Virtual Worlds* 25, 3-4 (2014), 201–211.
[72] Leonid Sigal and Michael J Black. 2006. Humaneva: Synchronized video and motion capture dataset for evaluation of articulated human motion. *Brown Univertsity TR* 120 (2006).
[73] Anselm Strauss and Juliet Corbin. 1994. Grounded Theory Methodology: An Overview. *Handbook of Qualitative Research* 273 (01 1994).
[74] Diana I. Tamir and Jason P. Mitchell. 2012. Disclosing information about the self is intrinsically rewarding. *Proceedings of the National Academy of Sciences* 109, 21 (2012), 8038–8043. https://doi.org/10.1073/pnas.1202129109 arXiv:https://www.pnas.org/content/109/21/8038.full.pdf
[75] S Trepte, L Reinecke, and K Behr. 2010. Avatar creation and video game enjoyment: effects of life-satisfaction, game competitiveness, and identification with the avatar world. In *Annual Conference of the International Communication Association, Suntec, Singapore*.
[76] Jan Van Looy, Cédric Courtois, Melanie De Vocht, and Lieven De Marez. 2012. Player identification in online games: Validation of a scale for measuring identification in MMOGs. *Media Psychology* 15, 2 (2012), 197–221.
[77] Martijn H. Vastenburg, Natalia Romero Herrera, Daan van Bel, and Pieter M.A. Desmet. 2011. PMRI: Development of a Pictorial Mood Reporting Instrument. In *CHI '11 Extended Abstracts on Human Factors in Computing Systems* (Vancouver, BC, Canada) *(CHI EA '11)*. ACM, New York, NY, USA, 2155–2160. https://doi.org/10.1145/1979742.1979933
[78] Timo von Marcard, Roberto Henschel, Michael J Black, Bodo Rosenhahn, and Gerard Pons-Moll. 2018. Recovering accurate 3d human pose in the wild using imus and a moving camera. In *Proceedings of the European Conference on Computer Vision (ECCV)*. 601–617.
[79] Cheng Yao Wang, Mose Sakashita, Upol Ehsan, Jingjin Li, and Andrea Stevenson Won. 2020. Again, Together: Socially Reliving Virtual Reality Experiences When Separated *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3313831.3376642
[80] Helen Wauck, Gale Lucas, Ari Shapiro, Andrew Feng, Jill Boberg, and Jonathan Gratch. 2018. Analyzing the effect of avatar self-similarity on men and women in a search and rescue game. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. 1–12.
[81] Yuanlu Xu, Song-Chun Zhu, and Tony Tung. 2019. Denserac: Joint 3d pose and shape estimation by dense render-and-compare. In *Proceedings of the IEEE International Conference on Computer Vision*. 7760–7770.
[82] Nick Yee and Jeremy Bailenson. 2007. The Proteus effect: The effect of transformed self-representation on behavior. *Human communication research* 33, 3 (2007), 271–290.

Received January 2021 ; revised April 2021 ; accepted May 2021