Measuring and Understanding Photo Sharing Experiences in Social Virtual Reality

Jie Li
Centrum Wiskunde & Informatica
Amsterdam, The Netherlands
jie.li@cwi.nl

Yiping Kong
Centrum Wiskunde & Informatica
Amsterdam, The Netherlands
yiping.kong@cwi.nl

Thomas Röggla
Centrum Wiskunde & Informatica
Amsterdam, The Netherlands
t.roggla@cwi.nl

Francesca De Simone
Centrum Wiskunde & Informatica
Amsterdam, The Netherlands
francesca.de.simone@cwi.nl

Swamy Ananthanarayan
University of Oldenburg
Oldenburg, Germany
s.ananthanarayan@uni-oldenburg.de

Huib de Ridder
Delft University of Technology
Delft, The Netherlands
h.deridder@tudelft.nl

Abdallah El Ali
Centrum Wiskunde & Informatica
Amsterdam, The Netherlands
abdallah.el.ali@cwi.nl

Pablo Cesar
Centrum Wiskunde & Informatica
Delft University of Technology
Amsterdam, The Netherlands
p.s.cesar@cwi.nl

ABSTRACT

Millions of photos are shared online daily, but the richness of interaction compared with face-to-face (F2F) sharing is still missing. While this may change with social Virtual Reality (socialVR), we still lack tools to measure such immersive and interactive experiences. In this paper, we investigate photo sharing experiences in immersive environments, focusing on socialVR. Running context mapping (N=10), an expert creative session (N=6), and an online experience clustering questionnaire (N=20), we develop and statistically evaluate a questionnaire to measure photo sharing experiences. We then ran a controlled, within-subject study (N=26 pairs) to compare photo sharing under F2F, Skype, and Facebook Spaces. Using interviews, audio analysis, and our questionnaire, we found that socialVR can closely approximate F2F sharing. We contribute empirical findings on the immersiveness differences between digital communication media, and propose a socialVR questionnaire that can in the future generalize beyond photo sharing.

CCS CONCEPTS

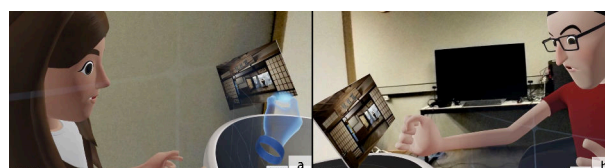• **Human-centered computing → Virtual reality**.

Figure 1: Photo sharing experience in Facebook Spaces© with remote partners (a) and (b) looking at the same photo.

KEYWORDS

Social, virtual reality, photo sharing, questionnaire, immersion, presence

1 INTRODUCTION

Ever since the introduction of film cameras in the late 19th century, photographs have been inherently social artifacts, which are captured and shared in social contexts [11]. Families and friends typically gathered in the living room, passing around printed photos or browsing albums [36] or they shared photos with each other via post [38]. Collocated (co-present) or remote photo sharing is a way of strengthening group bonds and interpersonal relationships [24]. Today, countless digital photos are uploaded to social media websites such as Instagram© [55] and Flickr© [36], or shared via

messaging apps such as WhatsApp© [10]. While online photos reach a wider audience [43], the asynchronous nature of the feedback affects the richness of face-to-face communication [41], especially between "strong ties" such as friends and family members [34, 35]. When users can schedule time together but are geographically apart, synchronous 2D video conferencing tools such as Skype© offer another viable alternative to meeting in person [25].

Recently, the wide availability of head-mounted displays (HMDs) for consumers makes immersive virtual reality (VR) systems a feasible alternative to audio communication and 2D video conferencing. With VR technology, people are able to "meet" in a shared, immersive virtual environment and interact with virtual representations of each other. Such environments with multiple users are denoted as collaborative or social VR (herefoth 'socialVR') [30]. The virtual representations in socialVR can map a user's movement onto an avatar (e.g., Facebook Spaces©[1] shown in Figure 1), use a 2D real-time user video as a virtual view of that user, or in the near future, use a highly detailed photo-realistic point cloud video [44]. While socialVR systems have high potential to allow several, physically separated users to communicate in a shared immersive virtual environment resembling face-to-face communication, it remains a challenge of how to measure such shared social experiences.

In this paper, we are interested in *measuring* and *understanding* how avatar-based socialVR and 2D video conferencing technologies can closely approximate face-to-face interactions surrounding photo sharing experiences. While many works have attempted to measure interaction experience, presence and immersion across real [32, 60] and virtual [26, 30, 53, 68] interactions, only recently has there been some work that addresses the socialVR medium in general (e.g., [30, 57, 67]), and no work that addresses socialVR photo sharing experiences in particular.

Furthermore, while social presence measurement tools can vary (e.g., subjective self-report measures [3] or non-verbal signals such as gestures [57]), we do not yet have a validated questionnaire that can capture the richness and social interaction nuances of photo-sharing in socialVR. Therefore, we ask: **(RQ1)** How can we adapt and develop a questionnaire to measure and understand immersiveness and social presence of socialVR photo sharing experiences? **(RQ2)** How does photo sharing in avatar-based socialVR compare with Skype-based communication and how well does it approximate face-to-face experiences? To answer our questions, we adopted a mixed-methods approach (shown in Figure 2) that combines a user-centric approach [47] and statistical techniques to develop an accurate and consistent questionnaire instrument (i.e., ensures test validity and reliability)
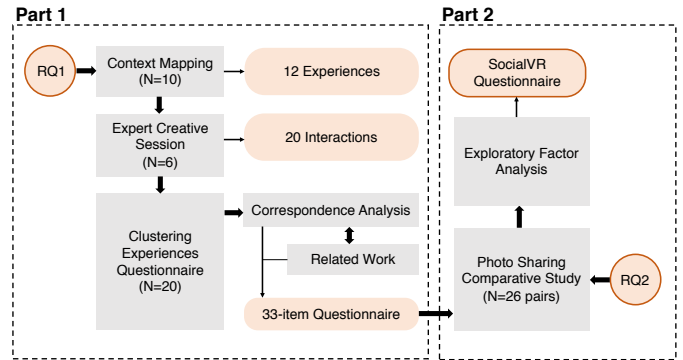
Figure 2: Methodological approach for constructing our SocialVR questionnaire.

for measuring socialVR photo sharing. The goal of our work is to mimic individuals sharing photos in the same physical space. This scenario lends itself particularly well to socialVR since it is a more immersive experience than for example Instagram or other asynchronous sharing systems. Similarly, for us to better understand co-present photo sharing, we needed to delve into social VR technology. This meant that we needed to understand more deeply notions of social presence and immersion, as these have an impact on co-present photo sharing.

Specifically, we ran context mapping sessions, an expert creative session, and an online experience clustering questionnaire, which resulted in three components of experiences that are important for photo sharing activities, namely quality of interaction (QoI), social meaning (SM) and presence/immersion (PI). Based on the three components and cross-checked with related work, we developed and statistically evaluated an initial questionnaire to measure socialVR photo sharing experiences. Next, we ran a controlled within-subject study (N=26 pairs) comparing photo sharing in face-to-face settings, Skype, and Facebook Spaces. Using exploratory factor analysis [12] on the collected data, we further validated our questionnaire. Then, using our questionnaire, interviews, and audio analysis techniques, we present empirical findings showing that Facebook Spaces can highly approximate face-to-face interactions for the social activity of photo sharing.

We make two primary contributions: **(1)** We develop and provide a questionnaire[2], designed through a user-centric process that was statistically evaluated to better understand social photo-sharing experiences across current, social communication media. **(2)** We provide empirical findings that help us understand how current socialVR technology affects

---

users' photo sharing experiences in comparison with face-to-face and 2D video conferencing. Below we start with a survey of related work.

2 RELATED WORK

Photo sharing has a vast and rich history in a variety of contexts from everyday consumer practices [45] to sharing images during times of disaster [39]. In this section, we focus on co-present photo sharing, followed by a review of communication behavior in various socialVR scenarios. Finally, we examine existing metrics for measuring presence/immersion experience in socialVR.

Co-present Photo Sharing

With the advent of camera phones, capturing photos is almost an everyday occurrence. Okabe et al. cite that one of the main reasons for taking photos has been for sharing memories with close friends and family [49]. By sharing photos, we not only express ourselves, but also nurture the relationships important to our lives [6]. Today, photos can be exchanged synchronously between people co-located in time and place or separated by distance. They can also be exchanged asynchronously by accessing a shared photo archive in a fixed location or by sending photos to others in a remote location [24]. Of these four options, co-present or face-to-face (F2F) photo sharing is regarded as the most enjoyable, and often seen as a way of reliving the experience with others [24]. Compared to digital photos uploaded to social media sites such as Instagram or Flickr where likes and comments are the currency of communication, co-present photo sharing allows the photographer to vividly communicate the experience through storytelling and emotion [66]. More importantly, digital photographs lack the tangibility and manipulability of physical photographs, an aspect considered important by many users [24].

Despite the lack of tangibility, digital photographs offer affordances (e.g., cost, immediacy, lifespan, flexibility) that are hard to ignore. Moreover, as van House et al. pointed out, co-present sharing is possible and important even under digitally-mediated worlds [62]. Paralanguage, body motion, proxemics, use of scent and artifacts are important even when sharing digital photos in F2F interactions [16]. Therefore, it is important to explore co-present photo sharing in technology-mediated communication systems. Particularly, we explore what related experiences are missing in the latest systems (e.g., VR, video calls) in comparison to F2F sharing.

Communication Behavior in Social VR

With the increasing prevalence of VR in the consumer market, distant users can now inhabit the same virtual environment and engage in a common activity. SocialVR has been employed in the automotive industry to enhance several design related activities [5]. In Zimmermann's survey, multiple automotive manufacturers mentioned that social VR assists in cross-discipline communication. For instance, it provides a platform for car design experts to create a scenario to communicate with managers who are not as intimately involved with the product [69]. Social scientists have also used VR to replicate and extend real-world experiments in a more controlled manner. VR in this case was helpful in creating stimuli that was impractical to achieve in the real world (e.g., large crowds, animals) [23]. Bombari et al. reviewed how VR has been used to study social interactions, with specific focus on the use of virtual humans as social interaction partners [8]. For instance, a few of the studies investigated how manipulating a virtual human's behavior influenced a participant's behavior [40]. Similarly, Ahn et al. [2] identified "embodied experience" in social VR with automated generation of VR scenes or objects based on user-provided data. This allowed users to see and hear as if they were experiencing someone else's point of view in the real world.

Another type of embodied VR is based on rendering users' movements onto an avatar in a virtual 3D environment. This supports nuanced nonverbal behavior alongside verbal communication. When comparing users in F2F, embodied VR with a full-body avatar, and no embodiment VR, researchers found that avatar embodiment provided a high level of social presence [57]; conversation patterns were very similar to F2F interactions. Although their work is similar to ours, Smith and Neff focus on communication patterns of same gender dyads for two different tasks rather than overall experience. Regardless, motion-controlled avatars with full representation of the avatar body lead to an increased sense of presence [30]. Consequently, we studied avatar-based socialVR for co-present photo sharing, especially since it requires a dynamic and improvisational interaction relationship.

Measuring Presence/Immersion in Social VR

The focus of measuring and understanding user experience (UX) is not only on utilitarian aspects (e.g., cognition and performance) of human-technology interactions, but also on affect, sensation, and the meaning as well as value of such interactions [37, 64]. Self-reported questionnaires have been the common methodology for evaluating UX. In the context of video conferencing and socialVR, metrics revolve around how we perceive and interpret the world around us. Particularly, the emphasis is on the ability to produce a sense of presence, or "being there", simulating F2F interactions. Metrics for evaluating presence and immersion have been developed and widely validated, such as the presence questionnaire by Witmer and Singer [68] and the Slater-Usoh-Steed questionnaire [56]. Furthermore, Jennett et al. [32] suggested, in their immersion questionnaire for gaming experience, to include

factors like lack of the awareness of time, involvement and a sense of being in the task environment. Indeed, there are already a variety of questionnaires to consider when evaluating socialVR. However, no single questionnaire is specifically suited for our task of photo sharing. Consequently, when developing our questionnaire we utilized items from the aforementioned questionnaires relevant to our work based on a user-centric approach.

3 PART 1: QUESTIONNAIRE DESIGN PROCESS

The goal of Part 1 (see Figure 2) was to identify typical experiences of photo sharing activities, and to develop a questionnaire for measuring socialVR photo sharing experiences later used in Part 2. Part 1 consisted of three steps: First, we started with two context mapping sessions [65] to gain insights into the current face-to-face (F2F) photo sharing, the experiences and phases of actions while sharing photos. Second, we ran a creative session with six experts in VR technology, aiming to better understand the social interactions around each phase of photo sharing. Third, with the identified experiences and the concluded social interactions identified by the experts, 20 UX designers and researchers were recruited to link the social interactions with the experiences. The relationship between the interactions and the experiences resulted in a further categorization of the experiences.

Context Mapping

Context mapping studies were conducted in a three-step approach: sensitizing participants with a designed booklet, group sessions, and data analysis and communication [65].

Participants. Ten participants (5 m, 5 f) were recruited. The age (M=25.1, SD=2.3) of the participants ranged between 22 and 31. Participants were recruited from X institute, and had varied nationalities. All had experience in face-to-face photo sharing with their friends and families.

Procedure. Before the session, participants were sensitized by filling in a paper-based booklet, and afterwards interviewed about their photo sharing experiences. The first part of the booklet requested participants to provide information about the forms of photo sharing, with whom they share, and locations and triggers. In the second part, participants were asked to describe / visualize their photo sharing experience on a timeline. Two group sessions were conducted. Each session invited five participants, and took approximately one hour. It started with peer discussions to compare their booklets (~10 min.), followed by presenting the differences and similarities of their responses (~10 min.). Later, participants were instructed to create a collage to visualize their ideal photo sharing experience (~20 min.), and thereafter present their visualizations to each other (~20 min.).
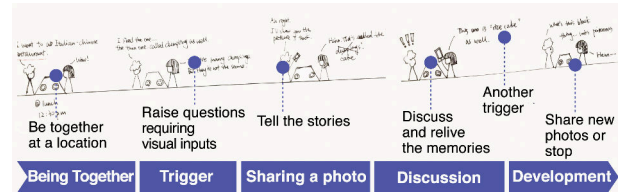
Figure 3: An example of experience timeline and five phases of actions during face-to-face photo sharing.

Data Collection and Analysis. All booklets and drawings were collected, and the audio-recorded interviews and sessions were transcribed. Three researchers read through all the booklets and transcriptions. Relevant contents were coded and interpreted, where coded contents were further divided into three categories: general information (e.g., what, with whom, where and how), 12 typical experiences, and phases of actions during F2F photo sharing.

Findings. Photo sharing activities were extensively discussed during the sessions. We found that digital photos are often displayed on mobile phones, which has no location restrictions in sharing (cf., [41]). Moreover, physical photos can be displayed in many forms, e.g., on a wall, frame or album. Sharing of physical photos was done preferably at home. Even though physical photos were the norm, participants indicated that digital sharing is the most convenient and frequent way of sharing, which typically happens between people who know each other. Sitting side-by-side was frequently mentioned as a common sitting position when sharing photos. Twelve typical F2F photo sharing experiences were identified during the sessions (summarized in Table 1 with representative quotes). Each participant created an experience timeline in their booklet to visualize one of their memorized photo sharing activities. The actions described on each timeline were coded and categorized by three researchers. In this way, the experience timelines converged into an overview of five phases of actions for sharing, which starts with getting together at the same location. Thereafter, people need a trigger to start the photo sharing (e.g., seeing a photo wall). The sharing process usually involves a shared focus on the photo content and storytelling, which is followed by discussions elicited by the photo. The process ends with a new round of photo sharing or changing topics. Figure 3 illustrates the five phases with an example of a timeline, where actions were coded with blue dots and interpreted with a phrase.

Expert Creative Session

Based on the mapped five phases of actions during F2F photo sharing, a group of six experts (E1-E6) (M=46.7, SD=14.4) were recruited to discuss the social interactions happening

Experience	Representative Quote	Participants
Sense of togetherness	*P8: I want to feel that I am together with my friend.*	P2, P5, P7, P8, P9, P10
Tangibility of photos	*P5: I want to touch and hold the photos.*	P2, P5, P7, P8, P9, P10
Engaging conversations	*P9: I want to have real-time responsive conversations.*	P5, P7, P9
Feel others' emotions	*P8: I want to understand my friend's emotions.*	P1, P2, P4, P7, P8, P9, P10
Feel understood	*P4: I want my friend to understand what I am talking about.*	P4, P6, P7, P8, P9, P10
Conversation triggers	*P1: The environment can inspire me to start conversations.*	P1, P3, P4, P5, P8, P9
Have fun	*P2: I want to enjoy the time spent with my friend.*	P1, P2, P3
Comfortable environment	*P6: I want to feel relaxed when sharing photos.*	P3, P4, P6, P8, P9
Show off	*P7: I want to show off my experiences.*	P6, P7
Create stronger bonds	*P6: The photo sharing enhances our relationship.*	P4, P6, P9, P10
Same focus	*P2: I want both of us to focus on the same thing.*	P2, P5, P6, P8, P9
Recall & recreate memories	*P3: I want to share my memories with my friends.*	P1, P2, P3, P4, P5, P6, P7, P8, P9, P10

Table 1: Typical experiences of face-to-face photo sharing.

during the sharing process. Expert profiles included three researchers, one software developer, one CEO and consultant, and an innovation manager, all of which work in the area of VR/AR technologies. Experts in VR were chosen as they are sensitive to the importance of immersive experiences.

Session Procedure. The creative session lasted approximately one hour. It began with a presentation by the session facilitator, introducing the context of F2F photo sharing and social interaction concepts (~5 min.). Experts were then instructed to brainstorm any possible social interactions that can happen between people (~15 min.). Next, the five phases of photo sharing (Figure 3) identified in the context mapping sessions were introduced. The experts were asked to cluster relevant social interactions they generated according to these five phases. They were also allowed to add new social interactions while clustering (~30 min.). The last step was to vote for the social interactions they considered important for F2F photo sharing (~10 min.).

Analysis and Findings. The five clusters generated according to the five phases were re-categorized and independently cross-checked by three researchers. The experts' votes were added up for the categories that emerged from the five phases. We used the categories that received at least three votes, which resulted in 20 categories. These are: way of speaking (13x), body gesture (10x), listening to atmospheric music (9x), eye contact (8x), pointing (8x), facial expressions (7x), hugs (5x), touch (5x), hand gestures (5x), laughter (5x), clothes / appearance (4x), gaze (4x), holding photos (4x), browsing albums together (4x), voice (4x), taking photos together (4x), playful activities (4x), drawing / marking (4x), asking questions (3x), mentioning others (3x).

## Clustering Experiences Questionnaire

Experience is shaped by the ways people interact and communicate in social relationships, known as social interactions

[58]. We also know interactions influence experience [14]. Forlizzi and Ford [22] proposed a model of user experience (UX) in terms of interactions, showing that experiences are momentary constructions that grow from interactions between people and their environment; here an "interaction-centered view" is valuable in understanding UX [21]. Based on this model, we further investigated how these 12 types of experience (identified in the context mapping sessions) can be clustered according to their relations with the 20 social interactions generated in the expert session. To identify such links, we designed a 12-item online questionnaire where respondents can map interactions to the 12 experiences identified earlier. Each item described an experience and provided a list of 20 interactions. Participants were requested to select the interaction(s) that can elicit the experience.

Participants. 20 respondents (14 f, 6 m) answered our questionnaire (M=25.6, SD=2.6). All respondents were either UX designers or researchers, the target group for this phase.

Correspondence Analysis. To identify the core components in the experience-interaction mappings, we ran a correspondence analysis [28]. This is an extension of principal component analysis (PCA) [33] that is well suited to explore relationships among categorical data (e.g., contingency tables). Like PCA, it provides a technique for summarizing and visualizing data in 2D plots. We did not use clustering algorithms (e.g., κ-means) since they are not directly applicable to categorical data, as the sample space for such data is discrete, and does not have a natural origin. We first tested the suitability of our data by running a Pearson's Chi-squared test (χ^2(209)=529.9), p<0.001). We then generated a symmetric plot with the experiences, which shows a global pattern within the data. Rows (interactions 1-20) are represented by gray points and columns (experiences 1-12) by red triangles. The distance between any row points or column points gives a measure of their similarity (or dissimilarity). Row
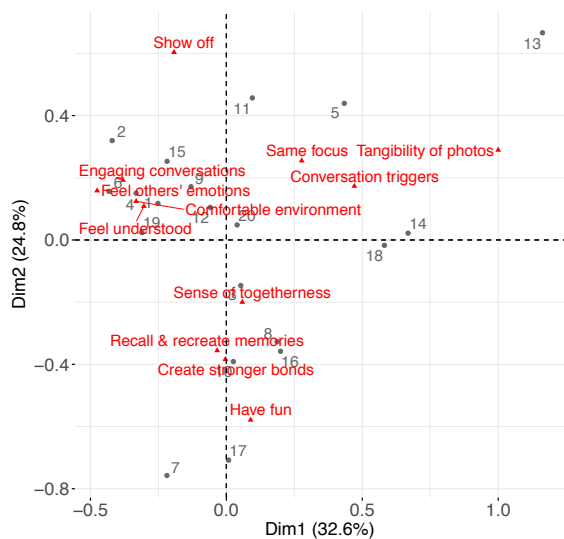
Figure 4: Correspondence analysis, showing three distinct experience components, along with the 20 (numerically displayed) interactions.

points with similar profiles are closer on the factor map, and similarly for column points.

Our correspondence analysis suggested three components of experience: *Quality of Interaction*, *Presence/Immersion*, and *Social Meaning*. From this analysis, we developed our photo sharing experience questionnaire. The questionnaire started with a self-report emotion tool developed by Vastenburg et al. [63]. This questionnaire was chosen (despite many existing emotion report questionnaires) as it is both rich (captures multiple categories) and an easy-to-use pictorial mood-reporting instrument. It is followed with questionnaire items that represent the three identified components of experience, all of which are cross-checked from existing presence questionnaires, and some subsequently adapted.

Component 1: Quality of Interaction (QoI). Four types of experience were included here: (1) feeling understood (2) engaging conversations (3) feeling others' emotions (4) comfortable environment. Except for (4), the other experiences are about how people communicate and interact with each other. Quality of interaction here is defined as the ability of the user to interact with the virtual world and to interact with other users in that virtual world [58, 59]. It assesses the quality of communication, mutual sensing of emotions, and naturalness between virtually represented users. Questionnaire items 1-11 in Table 2 were for measuring QoI, where items 1,2,8,9 were developed based on two experience types 1 and 3 in this component, items 3-7 adapted from Garau et al. [26], and items 10-11 adapted from Nilsson et al. [46].

Component 2: Presence/Immersion (PI). Three types of experience were included in this component: (1) conversation triggers (2) same focus (3) tangibility of photos. Witmer and Singer [68] defined immersion as a subjective and psychological state characterized by perceiving oneself to be involved in and interacting within a virtual environment. Questionnaire items 23-32 in Table 2 were for measuring PI, where item 23 was adapted from Slater et al. [56], items 24-25 from Schubert et al. [53], item 26 from Witmer & Singer [68], and items 27-32 from Jennett et al. [32].

Component 3: Social Meaning (SM). Four experience types were included in this component: (1) sense of togetherness (2) recall and recreate memories (3) create stronger bonds (4) have fun. Social meaning is defined as the experience of "being together", both mentally and physically. Questionnaire items 12-22 in Table 2 concerned measuring SM, where items 12-16 were adapted from Biocca et al. [7], items 19-21 adapted from van Bel et al. [61], item 22 adapted from Steen et al. [58], and items 17-18 developed based on experience types (2) and (3).

The resulting *32-item (+self-report emotions) questionnaire* questionnaire used in Part 2 is shown in Table 2, and found here: https://github.com/cwi-dis/socialVR_questionnaires

4 PART 2: PHOTO SHARING COMPARATIVE STUDY

In Part 2 (shown in Figure 2), we present a within-subject controlled user study compare photo sharing experiences in three conditions: face-to-face (F2F), Facebook Spaces (FBS), Skype (SKP). The resulting data is then: (a) used in an exploratory factor analysis (EFA) [12] to better understand the important factors in our questionnaire (b) provide empirical findings comparing photo sharing across study conditions.

Study Design
While many platforms have incorporated social capabilities in their VR experiences, FBS allows users to be immersed in a 360° photo-based virtual and shared space, using their self-customized avatar. The body and lip movements of the avatar are coordinated with user movements, and users can express a selection of facial emotions on the avatar using a controller. Users can also access their own photos, videos, and any media shared, which made FBS suitable for our photo sharing activity. While systems such as vTime[3] and High Fidelity[4] both have an in-built avatar-based socialVR community, where people can socialize with friends/family in 3D virtual destinations or in self-selected 360° photos, they have limitations in user control of facial emotions of

---

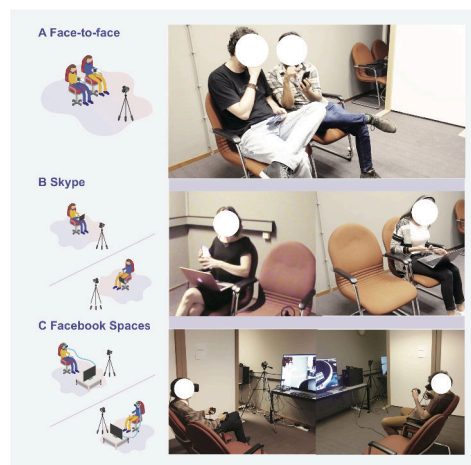the avatar. High Fidelity has no photo sharing function yet, and other socialVR platforms like SineSpace[5], Sansar[6] and AltSpaceVR[7] focus on inviting users to be creators and explorers of 3D virtual worlds, rather than supporting F2F photo sharing. While user avatars in these platforms can mimic users' body language, they are limited in what facial emotions can be expressed. Finally, we used FBS as it was one of the high-fidelity commercially available systems that required minimal setup. SKP was chosen as it is currently a standard in 2D video conferencing for friends/family to chat as if they are "face-to-face" [20], and importantly, can allow us to have a baseline by which to study current shared, social experiences.

Participants were asked to select three different photos on their smartphone for sharing purposes. Each participant shared one photo in each condition. The sequence of conditions was counterbalanced according to Latin Square design, and all sessions were video recorded (with consent). After each condition, participants filled in our developed 5-point Likert-scale questionnaire about the experience in that condition. A 5-point (and not 7-point) scale was chosen since with coarser measurements on lower sample sizes (typical of lab studies), there is lower variance. Furthermore, previous work has shown that 5- and 7-point scales are comparable if rescaling is performed [13]. Afterwards, participants filled a self-report emotion rating questionnaire [63]. Presence and immersion in a VR environment has been shown to play an important role in emotional reactions [15, 19], which led us to measure this subjectively (and analyze valence and arousal later automatically). A semi-structured interview was conducted when participants completed the three conditions and were sitting together. Interviews were audio recorded. Four main questions were asked during the interview: 1) Compared with F2F photo sharing, what do you think is missing in SKP and FBS? 2) How did you experience photo sharing activities in FBS and what else would you want to do in social VR? 3) To what extent are you satisfied with the virtual environment? 4) What for you is the future of immersive social communication media?

Participants

Twenty six (N=52) participant pairs (29 m, 23 f; M=27.6, SD=7.9) were recruited. All had diverse backgrounds and education levels. Recruiting criteria were: (a) Participant pairs for each experiment session should know each other and have a smartphone with at least three photos on it (b) Participants should be willing to share those photos with

[5]http://www.sine.space; retrieved: 20.09.2018
[6]https://www.sansar.com; retrieved: 20.09.2018
[7]https://altvr.com; retrieved: 20.09.2018

Figure 5: Illustrations of the three experimental conditions: (a) Face-to-face (b) Skype (c) FB Spaces.

their partner during sessions (c) Participants should not have visual or auditory problems to avoid motion sickness in VR.

Experiment Setup

Our experiment room was divided into two separate rooms by a movable wall. Both rooms have the same layout consisting of a pair of identical chairs, placed side by side. A computer with a 55" TV screen was placed in front of the chairs (Figure 5). For SKP, participants wore noise-cancelling headphones, and for FBS heard audio from the Oculus headset, so any potential proximity-related noise did not influence study participants. In the F2F condition, two participants were sitting together and showing each other photos on their smartphones and telling stories behind each photo. In the SKP condition, two participants were sitting in different rooms, where each was equipped with one laptop, and requested to share photos on their smartphones through SKP. In the FBS condition, photos were selected by participants and uploaded to the FBS system. Two participants were sitting in two different rooms, but they entered the same virtual space to share their photos (represented as physical photographs which are similar to the photo contents of a smartphone display). Although we could have uploaded photos to SKP, we opted for maintaining the same embodiment and self-representation (as in sharing photos in person) in all conditions. This also aligns with prior work that showed that users would spend more time in a video call if media sharing was easier [20]. Admittedly, there is a trade-off here with image quality and size but this approach preserves the tangible and physical nature of photograph sharing. Finally, to avoid repetition and keep it as natural as possible, participants shared different photos in the three conditions.

No.	Questionnaire items	Factor 1 (PI)	Factor 2 (SM)	Factor 3 (QoI)
1	"I was able to feel my partner's emotion during the photo sharing."			**0.61**
2	"I was sure that my partner often felt my emotion."			**0.67**
3	"It was easy for me to contribute to the conversation."	0.17	0.44	0.37
4	"The conversation seemed highly interactive."	0.36	0.26	0.33
5	"I could readily tell when my partner was listening to me."			**0.60**
6	"I found it difficult to keep track of the conversation."	-0.12	0.45	0.36
7	"I felt completely absorbed in the conversation."	0.33	0.44	0.18
8	"I could fully understand what my partner was talking about."		0.18	**0.71**
9	"I was sure that my partner understood what I was talking about."			**0.73**
10	"The experience of photo sharing seemed natural."	0.51		0.41
11	"The actions used to interact with my partner were natural."	**0.36**		0.24
12	"I often felt as if I was all alone during the photo sharing."		**0.62**	0.20
13	"I think my partner often felt alone during the photo sharing."		**0.62**	0.20
14	"I often felt that my partner and I were sitting together in the same space."	**0.82**		0.16
15	"I paid close attention to my partner."	0.14	0.12	**0.38**
16	"My partner was easily distracted when other things were going on around us."	-0.20	**0.32**	0.26
17	"I felt that the photo sharing enhanced our closeness."	**0.42**	0.21	
18	"Through the photo sharing, I managed to share my memories with my partner."	0.11	0.41	0.37
19	"I derived little satisfaction from photo sharing with my partner."	0.12	**0.56**	
20	"The photo sharing experience with my partner felt superficial."		**0.54**	0.18
21	"I really enjoyed the time spent with my partner."	0.18	0.43	0.29
22	"How emotionally close to your partner do you feel now?"	0.13	0.23	0.25
23	"I had a sense of being in the same space with my partner."	**0.92**		
24	"Somehow I felt that the same space was surrounding me and my partner."	**0.87**	0.12	-0.15
25	"I had a sense of interacting with my partner in the same space, rather than doing it through a system."	**0.88**		
26	"My photo sharing experience seemed as if it was a face-to-face sharing."	**0.80**	-0.22	0.27
27	"I did not notice what was happening around me during the photo sharing."*	**0.52**	0.30	-0.12
28	"I felt detached from the world around me during the photo sharing."*	**0.71**	0.20	-0.20
29	"At the time, I was totally focusing on photo sharing."	0.36	0.38	
30	"Everyday thoughts and concerns were still very much on my mind."		**0.69**	-0.16
31	"It felt like the photo sharing took shorter time than it really was."	0.25	**0.31**	
32	"When sharing the photos, time appeared to go by very slowly."	-0.10	**0.54**	
	SS loadings	5.67	3.83	3.65
	Proportion Variance	0.18	0.12	0.11
	Cumulative Variance	0.18	0.29	0.41

Factor loadings of 0.3 and above and without cross-loadings of >0.3 are marked in bold.

Table 2: Exploratory factor analysis (EFA) applied to our questionnaire items, where questions in bold indicate that these items are kept for the final analysis.

5 RESULTS

Exploratory Factor Analysis

We ran an exploratory factor analysis (EFA) [12] to better understand the important factors in our questionnaire. EFA is a statistical technique within factor analysis commonly used for scale development involving categorical and ordinal data, and serves to identify a set of latent constructs underlying a battery of measured variables [18, 48]. Given that our focus was on evaluating SKP and FBS, and that they contained the complete list of questions, we ran our analysis only on data from these two system evaluations. Since Bartlett's Sphericity test was significant ($\chi^2(2,496)$ = 2207.187, p<0.001) and

Kaiser-Meyer-Olkin was greater than 0.5 (KMO=0.85), our data allowed for EFA. Given our earlier correspondence analysis that showed a grouping of three factors, we tested our model fit based on three factors corresponding to each set of questionnaire items. Furthermore, since we assumed that factors would be related, we used oblique rotation ('oblimin') along with standard principal axes factoring. Standardized loadings are shown in Table 2.

To ensure the factors are meaningful and redundancies eliminated (removing collinearity effects), we only took items with factor loadings of 0.3 and above, and with cross-loadings not less than 0.2 across factors. The cumulative explained variance of the three factors is 41%. The 24 questionnaire

items in bold were used for our evaluation of the three conditions (F2F, FBS, and SKP) along the identified concepts: *Quality of Interaction (QoI)*, *Social Meaning (SM)*, and *Presence/Immersion (PI)*. We furthermore tested each set of items for internal reliability by measuring Cronbach's alpha, and our final item sets show high reliability coefficients: F2F QoI (α=0.8), F2F SM (α=0.89), F2F PI (α=0.74), FBS QoI (α=0.79), FBS SM (α=0.83), FBS PI (α=0.76), SKP QoI (α=0.78), SKP SM (α=0.79), SKP PI (α=0.75). Our final 24-item *Social Virtual Reality Photo Sharing Experiences Questionnaire* is shown in Table 2 with the included items in bold, and found on Github: https://github.com/cwi-dis/socialVR_questionnaires

Questionnaire Response Analysis

We consider the effects of the three factors (F2F, FBS, SKP) on each 5-point likert-scale measure: Quality of Interaction (QoI), Social Meaning (SM), and Presence/Immersion (PI).

Quality of Interaction (QoI). The sum of scores means and standard deviations for the QoI questions (6 items) for each tested photo sharing condition are: F2F=25.9(4.3), FBS=22.6(4.3), SKP=21.4(4). The sum of scores are compared in Figure 6a. 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. A Shapiro-Wilk-Test showed that our data is not normally distributed (p<0.001). As we compare three matched groups within subjects, we directly performed a Friedman rank sum test. Here we found a significant effect of photo sharing conditions on QoI (χ^2(2)=39.1), p<0.001). A post-hoc test using Mann-Whitney tests with Bonferroni correction showed significant differences between F2F and FB Spaces (p<0.001, r = 0.41), between F2F and SKP (p<0.001, r = 0.54), but not between FBS and SKP (p=0.07). These results indicate that with respect to QoI, F2F photo sharing was perceived to be better than both FBS and SKP.

Presence/Immersion (PI). Since F2F did not involve interacting with a photo sharing system and required four PI item omissions from our questionnaire, a comparative analysis is not applicable here[8]. The sum of scores means and standard deviations for the PI questions (9 items) for each tested photo sharing system interaction are: FBS=34.9(6), SKP=21.7(5.9). The sum of scores are compared in Figure 6b. A Shapiro-Wilk-Test showed that our data is not normally distributed (p<0.001). As we compare two matched groups within subjects, we directly performed a Wilcoxon rank sum test. Here we found a significant effect of photo sharing conditions on PI (W=2511, p<0.001, η_p^2=0.74). These results indicate that with respect to PI, FBS was perceived to be more immersive and result in higher feelings of presence than SKP.

---

[8]As a sanity check, if we assume perfect scores for the missing items, given that F2F is our gold standard, the mean and standard deviation would be 38.1(3.3).

Social Meaning (SM). The sum of scores means and standard deviations for SM (9 items) for each tested photo sharing condition are: F2F=34.2(6.2), FBS=35.2(5.7), SKP=30.4(5.4). The sum of scores are compared in Figure 6c. A Shapiro-Wilk-Test showed that our data is not normally distributed (p < 0.001). As we compare three matched groups within subjects, we directly performed a Friedman rank sum test. Here we found a significant effect of photo sharing conditions on QoI (χ^2(2)=27.2), p<0.001). A post-hoc test using Mann-Whitney tests with Bonferroni correction did not show significant differences between F2F and FBS (p=0.55), however did between F2F and SKP (p<0.001, r = 0.37) and between FBS and SKP (p<0.001, r = 0.43). These results indicate that with respect to QoI, FBS photo sharing was comparable with F2F interactions, and significantly different than with SKP.

Emotion Ratings

We plotted the emotion ratings on a radar chart (Figure 6d), which shows FBS was perceived to be more exciting and cheerful than F2F and SKP. The sum of scores means and standard deviations (after subtracting the negative ratings) for emotions (N=8) under each photo sharing condition are: F2F=17.6(8.9), FBS=16.8(8.7), SKP=15.6(9.3). These are visually compared in Figure 6d. A Shapiro-Wilk-Test showed that our data is not normally distributed (p < 0.001). As we compare three matched groups within subjects, we performed a Friedman rank sum test. We found a significant effect of photo sharing condition on emotion ratings (χ^2(2)=8.74), p<0.05). However, running post-hoc Mann-Whitney tests with Bonferroni correction did not show any significant interaction effects.

Audio Feature Analysis

We further analyzed the audio in each video to test for differences in valence, arousal, and overall speech duration across each condition. We used PyAudioAnalysis [27], an open-source Python library for audio signal analysis. While we had multiple recording microphones for SKP and FBS, we analyzed the audio data from the microphone placed equidistant between each participant.

Valence and Arousal: We looked at whether valence and arousal from the recorded audio streams differed across photo sharing conditions. These are computed using PyAudioAnalysis library's SVM regression training in order to map audio features to one or more supervised variables. We used the built-in training examples (detailed in [27]). Our results are summarized in Table 3. For valence, we find that F2F and FBS have higher detected valence values than SKP, however this difference is not statistically significant using a paired t-test. With respect to arousal, we found that photo sharing interactions with all systems were slightly

(a) Quality of Interaction (QoI). **(b) Presence/Immersion (PI).** **(c) Social Meaning (SM).** **(d) Reported emotions.**
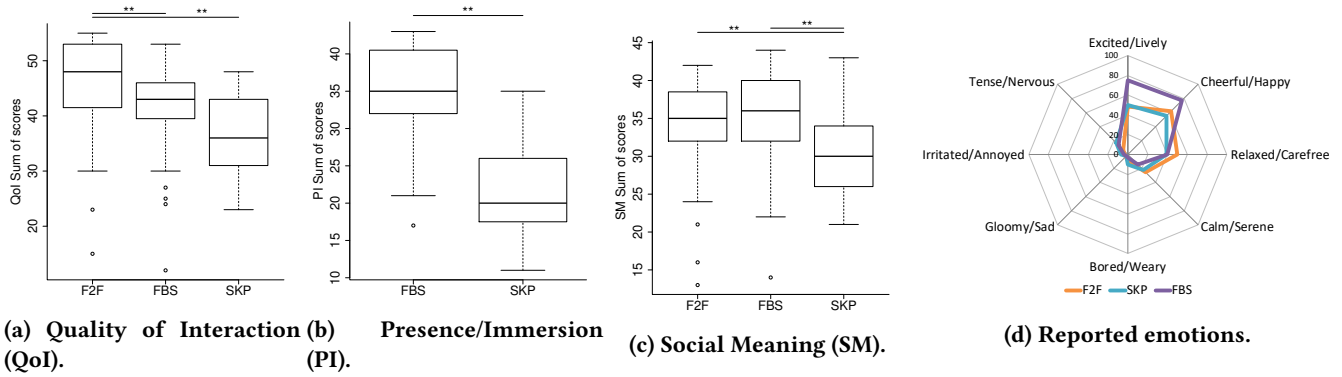
Figure 6: (a)-(c) Sum of scores boxplots for across photo sharing conditions for face-to-face (F2F), FB Spaces (FBS), Skype (SKP). (d) Self-reported emotion ratings for each condition. ** = p<.001

	Valence	Arousal
F2F	0.63	0.29
FB Spaces	0.63	0.32
Skype	0.56	0.36

Table 3: Mean detected valence and arousal of audio streams across conditions.

arousing, with SKP showing the highest value. Here, only the difference between F2F and SKP was statistically significant (t(25)=-2.42, p<0.05, Cohen's d=0.58), where a t-test was appropriate here given the data was normally distributed under a Shapiro-Wilk test (p=0.60). We further tested whether these values correlate (using Spearman's ρ) with the reported emotion ratings, and found that detected F2F valence (0.77) and arousal (0.62) exhibit high correlations with self-reports, that detected SKP valence (0.43) and arousal (0.38) exhibit medium correlations with self-reports, and also medium correlations for FBS for both valence (0.50) and arousal (0.55). These findings highlight that automatic detection of valence and arousal from social interaction audio streams can reasonably correlate with self-report data.

Speech Duration and Turn Taking: We further analyzed speech durations across each conditions, and found that F2F resulted in highest amount of speech (83.8%), followed by FBS (78.1%) and SKP (75%). Since this data was not normally distributed (P<0.01), we used a Friedman rank sum test, however we did not find any significant differences (p=0.07). Furthermore, we looked at the mean (and standard deviations) conversational turns per minute across conditions: F2F: 4.76(2.3), FBS: 5.73(2.5), SKP: 4.94(2.8). Since our data was not normally distributed for all turn taking across conditions (P<0.05), we again ran a Friedman rank sum test, however did not find significant differences across conditions (p=0.17).

Interview Analysis

Audio recordings of the semi-structured interviews were transcribed and coded by two researchers, following an open coding approach [52]. Following this approach, we only used the themes that emerged from both independent coders, so we only considered themes with full agreement. From the coded transcripts, four main themes emerged, which we discuss below. Twenty-six pairs of participants are labeled P1A(B)-P26A(B).

Limitations of screen-based communication. Half the participants (52%) expressed concerns over using SKP for photo sharing, where they felt separated by a monitor screen. They also felt distracted by the environment and became less focused on the conversation (P23B: *"Skype is kind of curtain between you two."*). Furthermore, the mismatch between the camera and the screen of the laptop did not allow participants to have eye contact (P9B: *"The position of the camera makes the eye contact impossible in Skype."*). Moreover, while this is a broader issue where internet connections can be affected by bandwidth issues, some participants (21%) complained that the occasional delay in the SKP audio stream influenced the fluency of the conversations.

Avatar-based embodiment in social VR. For FBS, a few participants (13%) complained that the avatar appearance is too detached from reality (P1B: *"It is difficult to link avatars with human beings. I will put more attention on the voice, rather than the avatar face."*). Participants (29%) did not feel that they actually saw each other (P1A: *"Avatar reminds you that you are still in VR."*). Another missing aspect mentioned (35%) is the limited ability to show facial expressions on their avatars via the Oculus Swift controllers, where these sometimes felt unnatural and restrictive in expressiveness (P25A: *"In VR, you cannot express your emotions because you only have a few options, and you need to control them with buttons."*).

Furthermore, timely choosing as emotion was found to be difficult (P2A: *"It is difficult to think about the emotions I need to show on my avatar, when I'm pointing to the photo or telling stories."*). Instead, participants relied on extracting emotional cues from voice (P22B: *"It's a bit difficult to show emotions, but you can still hear their [partner's] voice and interpret it."*).

Social VR immersiveness and novelty effects. Compared with SKP, participants felt physically and emotionally closer using FBS (P18B: *"In VR, you feel much closer than in Skype, like staying in the same room."*). An interesting aspect raised here is that the more detached from the real world participants felt, the more they could focus on each other (P13A: *"I felt we were in the same space, and focusing on the same activity"*). Throughout, the "wow" effect of FBS was raised (38%). Participants tended to feel more excited and happier because of the new technology (P9A: *"VR experience was very exciting!"*), however, this came at a technological habituation cost (P24A: *"I guess, with time, I will get used to it."*).

Beyond photo sharing in social VR. Participants believed that social VR can bring new forms of social interactions (25%). Using such technology, participants would like to do activities that were not possible in the real world. For instance, the most frequently mentioned activities included gaming (31%), collaborating in 3D spaces (25%), family or friend gatherings (21%), and exploring the world (19%). Participants also suggested that social VR should explore novel social activities, aside from everyday interactions like photo sharing (P16B: *"We need some novel interactive approaches to live another person's life, for instance."*).

Future of social VR. Most participants (87%) were satisfied with the socialVR environment. Even though the resolution of the background 360° image was not ideal, they felt immersed (P3B: *"The borderless view made me it feel very real."*). They mentioned that immersion could be improved by using high resolution images and by reducing the weight of the HMD (P4B: *"My HMD is not comfortable. It is heavy and tight, which reminds me that I am in VR."*). Importantly, participants (33%) suggested that future socialVR platforms for meeting strangers and close friends should be separated due to safety and privacy concerns. Indeed, this echoes previous work that showed users' privacy concern when sharing photos in collocated [41] and public settings [31].

6 DISCUSSION
Limitations
Our first limitation was that we restricted ourselves to avatar-based, and not photo-realistic socialVR, even though realistic avatars are rated more human-like [4, 50]. While this is one

instance of socialVR, this was the current commercially available state-of-the-art suitable for photo sharing. Second, in our lab study, we saw that FBS produced novelty effects in some participants, which is not indicative of longer-term usage, and can result in potential technological habituation. Nevertheless, given our participant sample size and the short-term nature of current VR usage, this should not undermine participant responses. Third, while we ensured construct, internal, and face validity of our questionnaire, we only took measures to ensure external validity through one controlled study. To increase external validity, researchers and practitioners would need to test this socialVR questionnaire across multiple experiments. Fourth, as in recent prior work [57], our focus was also on a single activity (photo sharing), and explored it fully; however, we are aware perceptions might differ across other contexts and activities.

Measuring Social VR Interaction and Immersiveness
An overarching goal of this work is to equip researchers with tools to measure photo sharing interaction and immersiveness in this new socialVR communication medium **(RQ1)**. It is important here to ask why we need another questionnaire, given that many previous works have also measure interaction experience, presence and immersion across real [32, 60] and virtual [26, 30, 53, 68] interactions. Based on our literature review, we find that socialVR is a new medium where photo sharing activities have not been studied. Here, we took existing questionnaires along with our user-centric approach [47] as a starting point, where we ensured that designers, researchers, and importantly end-users are key contributors to defining experiences relevant to photo sharing. While we initially adopt a qualitative approach to find *what* to measure, our work advances this using quantitative methods that helped make sense of the sometimes messy qualitative data that can hinder the development of a Likert-scale questionnaire. This helps answer whether what we measured is, based on statistical data reduction techniques, what is useful to measure in such socialVR contexts.

Furthermore, by drawing on audio analysis techniques, we provided early steps to automatically capture aspects of users' experiences (valence and arousal, turn taking, speaking duration) during socialVR interactions. Contrary to prior work (e.g., [57]), our goal here was not to understand communication signals per se, but rather eventually provide a set of tools to measure immersiveness for the specific activity of photo sharing while ensuring capture of rich subjective feedback from users. In this regard, even with automatic measurement (e.g., user motion analysis) and witnessing correlations between measures (e.g., audio-based valence/arousal and subjective emotion ratings), we believe it is important to have a representative questionnaire that captures the richness of social experiences in VR worlds.

Social VR Photo Sharing Approximates F2F

Social avatar-based VR is perceived to be still lacking compared with F2F, but it raises the question whether it will replace or augment F2F activities such as photo sharing? Based on our quantitative findings, we find that F2F scored highest on QoI, with FBS and SKP both scoring lower. Recent work showed that computer-generated virtual avatars require full mapping to human motion [30], and that believable body motion is more important than head motion [42]. This may be unsurprising given the visceral nature of shared reality experiences [17]. However, it was interesting that with respect to SM, both F2F and FBS were perceived similarly and scored higher than SKP. This finding was echoed in our interviews in that participants found FBS to provide a more physical and closer connection to a partner than SKP. This could be a side effect of being in a completely virtual and immersive world, where among virtual objects users found solace in establishing a meaningful connection with their human partner, even if represented virtually as a cartoon-like avatar. Indeed, if we look at PI comparing FBS and SKP, we see that FBS scored significantly higher than SKP.

Digging deeper, we found that FBS was rated as an exciting and happy experience, and corroborated by our speech valence analysis. Furthermore, while F2F on average resulted in more speech, turn taking across all conditions was similar. Together, our analysis underscores the potential of socialVR to closely approximate F2F (**RQ2**), at least for current social practices of collocated F2F photo sharing. This supports Smith & Neff's [57] findings, who found that embodied VR provides users with high levels of social presence and conversation patterns that are highly similar to F2F interactions.

Beyond Real-world Photo Sharing in Social VR

During interviews, a quarter of participants stated they would like to perform activities in socialVR that are not possible in the real world. While exploring imaginary activities was not in our scope, it raises a question about the role of socialVR: are we more concerned with adapting our real-world social activities to socialVR as is, or do we want to infuse our social environment with imaginary elements (e.g., sharing photos together on a virtual mountain)? While such questions may seem far away from current socialVR technology, it highlights not only the role of embodiment in activities such as photo sharing, but also what type of embodiment we assume [67]. For example, Schwind et al. [54] showed that women perceive lower levels of presence while using male avatar hands. We can speculate whether our future communication tools (e.g., video conferencing) should simulate F2F interactions as realistically as possible, and whether the underlying assumption that our current "gold standard" of comparing against F2F interactions will be the baseline of the future.

Measuring Sharing Activities across Contexts

While our work and questionnaires were geared towards photo sharing in socialVR, our work can be adapted to other social activities that are grounded in social interactions we draw from experience. From our correspondence analysis, we arrived at three experience components: (1) *Quality of Interaction*, defined as the ability of the user to interact with the virtual world and to interact with other users in that world [58, 59] (2) *Presence/Immersion*, defined as a subjective and psychological state characterized by perceiving oneself to be involved in and interacting within a virtual environment [68] (3) *Social Meaning*, the social aspect of presence, defined as the experience of "being together", both mentally and physically [7]. All of these components serve as building blocks of F2F experiences in socialVR spaces.

In this respect, sharing photos can be viewed as one kind of sharing, and our questionnaires can be adapted to address other F2F activities. These can be slight variants of photo sharing such as sharing artwork in the virtual cultural heritage domain [1], but also extends to other contexts such as sharing and watching videos together [29], sharing knowledge (e.g., teaching children by doing [51]), or even collocated F2F games (e.g., board games [9]). Due to the generality of our surfaced experience components and scale constucts, researchers and practitioners can easily adapt our contributed questionnaires for other sharing tasks.

7 FUTURE WORK AND CONCLUSION

While we touched on automatic capture and analysis of socialVR behavior, we are currently exploring other interaction signals, including head pose, jitter/movement, and affect signals. With new advances in socialVR, we set out to investigate photo sharing experiences. We followed a user-centric approach (context mapping, expert session, experience clustering questionnaire), in order to adapt, develop, and statistically evaluate a questionnaire that can be used as a standard for measuring socialVR photo sharing experiences. As a first step towards external validity, we ran a controlled user study to compare photo sharing experiences under F2F, SKP, and FBS. Based on our interviews, audio analysis, and our developed questionnaire, we found that socialVR is capable of closely approximating F2F sharing. While our work raised the possibility that future socialVR systems do not have to imitate reality-based F2F experiences, our questionnaire contribution and our empirical findings concerning photo sharing experiences can provide researchers and designers with a tool to measure such reality-grounded activities.

8 ACKNOWLEDGMENTS

REFERENCES

[1] A. C. Addison. 2000. Emerging trends in virtual heritage. *IEEE Multi-Media* 7, 2 (April 2000), 22–25. https://doi.org/10.1109/93.848421

[2] Sun Joo (Grace) Ahn, Amanda Minh Tran Le, and Jeremy Bailenson. 2013. The Effect of Embodied Experiences on Self-Other Merging, Attitude, and Helping Behavior. *Media Psychology* 16, 1 (2013), 7–38. https://doi.org/10.1080/15213269.2012.755877

[3] William Albert and Thomas Tullis. 2013. *Measuring the user experience: collecting, analyzing, and presenting usability metrics.* Newnes.

[4] Jeremy N. Bailenson, Nick Yee, Dan Merget, and Ralph Schroeder. 2006. The Effect of Behavioral Realism and Form Realism of Real-Time Avatar Faces on Verbal Disclosure, Nonverbal Disclosure, Emotion Recognition, and Copresence in Dyadic Interaction. *Presence: Teleoper. Virtual Environ.* 15, 4 (Aug. 2006), 359–372. https://doi.org/10.1162/pres.15.4.359

[5] Leif P. Berg and Judy M. Vance. 2017. Industry Use of Virtual Reality in Product Design and Manufacturing: A Survey. *Virtual Real.* 21, 1 (March 2017), 1–17. https://doi.org/10.1007/s10055-016-0293-9

[6] Margit Biemans, Betsy van Dijk, Pavan Dadlani, and Aart van Halteren. 2009. Let's stay in touch: sharing photos for restoring social connectedness between rehabilitants, friends and family. In *Proc. SIGACCESS '09.* ACM, 179–186.

[7] Frank Biocca, Chad Harms, and Jenn Gregg. 2001. The networked minds measure of social presence: Pilot test of the factor structure and concurrent validity. In *4th annual international workshop on presence, Philadelphia, PA.* 1–9.

[8] Dario Bombari, Marianne Schmid Mast, Elena Canadas, and Manuel Bachmann. 2015. Studying social interactions through immersive virtual environment technology: virtues, pitfalls, and future challenges. *Frontiers in psychology* 6 (2015), 869.

[9] Llogari Casas, Loïc Ciccone, Gökçen Çimen, Pablo Wiedemann, Matthias Fauconneau, Robert W. Sumner, and Kenny Mitchell. 2018. Multi-reality Games: An Experience Across the Entire Reality-virtuality Continuum. In *Proc. VRCAI '18.* ACM, New York, NY, USA, Article 18, 4 pages. https://doi.org/10.1145/3284398.3284411

[10] Ying-Yu Chen, Frank Bentley, Christian Holz, and Cheng Xu. 2015. Sharing (and Discussing) the Moment: The Conversations That Occur Around Shared Mobile Media. In *Proc. MobileHCI '15.* ACM, New York, NY, USA, 264–273. https://doi.org/10.1145/2785830.2785868

[11] James Clawson, Amy Voida, Nirmal Patel, and Kent Lyons. 2008. Mobiphos: A Collocated-synchronous Mobile Photo Sharing Application. In *Proc. MobileHCI '08.* ACM, New York, NY, USA, 187–195. https://doi.org/10.1145/1409240.1409261

[12] Anna B Costello and Jason W Osborne. 2005. Best practices in exploratory factor analysis. *Practical Assess., Research & Eval.* 10, 7 (2005), 1–9.

[13] Dr. John Dawes. 2008. Do Data Characteristics Change According to the Number of Scale Points Used? An Experiment Using 5-Point, 7-Point and 10-Point Scales. *International Journal of Market Research* 50, 1 (2008), 61–104. https://doi.org/10.1177/147078530805000106

[14] Pieter MA Desmet and Paul Hekkert. 2007. Framework of product experience. *International journal of design* 1, 1 (2007), 57–66.

[15] Julia Diemer, Georg W. Alpers, Henrik M. Peperkorn, Youssef Shiban, and Andreas Mühlberger. 2015. The impact of perception and presence on emotional reactions: a review of research in virtual reality. *Frontiers in Psychology* 6, 26. https://doi.org/10.3389/fpsyg.2015.00026

[16] Starkey Duncan and Donald W Fiske. 2015. *Face-to-face interaction: Research, methods, and theory.* Routledge.

[17] Gerald Echterhoff, E. Tory Higgins, and John M. Levine. 2009. Shared Reality: Experiencing Commonality with others' Inner States about the World. *Perspectives on Psychological Science* 4, 5 (2009), 496–521.

https://doi.org/10.1111/j.1745-6924.2009.01161.x PMID: 26162223.

[18] Leandre R Fabrigar, Duane T Wegener, Robert C MacCallum, and Erin J Strahan. 1999. Evaluating the use of exploratory factor analysis in psychological research. *Psychological methods* 4, 3 (1999), 272.

[19] Anna Felnhofer, Oswald D. Kothgassner, Mareike Schmidt, Anna-Katharina Heinzle, Leon Beutl, Helmut Hlavacs, and Ilse Kryspin-Exner. 2015. Is virtual reality emotionally arousing? Investigating five emotion inducing virtual park scenarios. *International Journal of Human-Computer Studies* 82 (2015), 48 – 56. https://doi.org/10.1016/j.ijhcs.2015.05.004

[20] Azadeh Forghani, Gina Venolia, and Kori Inkpen. 2014. Media2Gether: Sharing Media During a Call. In *Proc. GROUP '14.* ACM, New York, NY, USA, 142–151. https://doi.org/10.1145/2660398.2660417

[21] Jodi Forlizzi and Katja Battarbee. 2004. Understanding experience in interactive systems. In *Proc. DIS '04.* ACM, 261–268.

[22] Jodi Forlizzi and Shannon Ford. 2000. The building blocks of experience: an early framework for interaction designers. In *Proceedings of the 3rd conference on Designing interactive systems: processes, practices, methods, and techniques.* ACM, 419–423.

[23] Jesse Fox, Dylan Arena, and Jeremy N. Bailenson. 2009. Virtual reality: A survival guide for the social scientist. *Journal of Media Psychology* 21, 3 (2009), 95–113.

[24] David Frohlich, Allan Kuchinsky, Celine Pering, Abbe Don, and Steven Ariss. 2002. Requirements for Photoware. In *Proc. CSCW '02.* ACM, New York, NY, USA, 166–175. https://doi.org/10.1145/587078.587102

[25] Susan R. Fussell and Leslie D. Setlock. 2014. Computer-Mediated Communication. *The Oxford Handbook of Language and Social Psychology* (Sept. 2014). https://doi.org/10.1093/oxfordhb/9780199838639.013.020

[26] Maia Garau, Mel Slater, Vinoba Vinayagamoorthy, Andrea Brogni, Anthony Steed, and M. Angela Sasse. 2003. The impact of avatar realism and eye gaze control on perceived quality of communication in a shared immersive virtual environment. In *In Proceedings of the SIGCHI conference on Human factors in computing systems.* 529–536.

[27] Theodoros Giannakopoulos. 2015. pyAudioAnalysis: An Open-Source Python Library for Audio Signal Analysis. *PloS one* 10, 12 (2015).

[28] Michael Greenacre. 2017. *Correspondence analysis in practice.* Chapman and Hall/CRC.

[29] Simon Gunkel, Hans Stokking, Martin Prins, Omar Niamut, Ernestasia Siahaan, and Pablo Cesar. 2018. Experiencing Virtual Reality Together: Social VR Use Case Study. In *Proc. TVX '18.* ACM, New York, NY, USA, 233–238. https://doi.org/10.1145/3210825.3213566

[30] Paul Heidicker, Eike Langbehn, and Frank Steinicke. 2017. Influence of avatar appearance on presence in social VR. In *3D User Interfaces (3DUI), 2017 IEEE Symposium on.* IEEE, 233–234.

[31] Jussi Holopainen, Andrés Lucero, Hannamari Saarenpää, Timo Nummenmaa, Abdallah El Ali, and Tero Jokela. 2011. Social and Privacy Aspects of a System for Collaborative Public Expression. In *Proc. ACE '11.* ACM, New York, NY, USA, Article 23, 8 pages. https://doi.org/10.1145/2071423.2071452

[32] Charlene Jennett, Anna L. Cox, Paul Cairns, Samira Dhoparee, Andrew Epps, Tim Tijs, and Alison Walton. 2008. Measuring and defining the experience of immersion in games. *International journal of human-computer studies* 66, 9 (2008), 641–661.

[33] Ian Jolliffe. 2011. Principal component analysis. In *International encyclopedia of statistical science.* Springer, 1094–1096.

[34] Daniel Kahneman and Patrick Egan. 2011. *Thinking, fast and slow.* Vol. 1. Farrar, Straus and Giroux New York.

[35] Joke Kort, Harold T Nefs, Charlie Gullström, Tjerk de Greef, and Peter Parnes. 2013. Connected media and presence. In *Proceedings of the 2nd international workshop on Socially-aware multimedia.* ACM, 43–48.

[36] Christian Kray, Michael Rohs, Jonathan Hook, and Sven Kratz. 2009. Bridging the gap between the Kodak and the Flickr generations: A

novel interaction technique for collocated photo sharing. *International Journal of Human-Computer Studies* 67, 12 (2009), 1060–1072.

[37] Effie Lai-Chong Law, Virpi Roto, Marc Hassenzahl, Arnold POS Vermeeren, and Joke Kort. 2009. Understanding, scoping and defining user experience: a survey approach. In *Proceedings of the SIGCHI conference on human factors in computing systems*. ACM, 719–728.

[38] Siân E. Lindley, Abigail C. Durrant, David S. Kirk, and Alex S. Taylor. 2008. Collocated Social Practices Surrounding Photos. In *Proc. CHI '08 EA (CHI EA '08)*. ACM, New York, NY, USA, 3921–3924. https://doi.org/10.1145/1358628.1358957

[39] Sophia Liu, Leysia Palen, Jeannette Sutton, Amanda Hughes, and Sarah Vieweg. 2008. In Search of the Bigger Picture: The Emergent Role of On-Line Photo-Sharing in Times of Disaster. In *Proceedings of the Information Systems for Crisis Response and Management Conference (ISCRAM 2008)*.

[40] Joan Llobera, Bernhard Spanlang, Giulio Ruffini, and Mel Slater. 2010. Proxemics with multiple dynamic characters in an immersive virtual environment. *ACM Transactions on Applied Perception* 8, 1 (2010), 3.

[41] Andrés Lucero, Jussi Holopainen, and Tero Jokela. 2011. Pass-them-around: Collaborative Use of Mobile Phones for Photo Sharing. In *Proc. CHI '11*. ACM, New York, NY, USA, 1787–1796. https://doi.org/10.1145/1978942.1979201

[42] Elhassan Makled, Yomna Abdelrahman, Noha Mokhtar, Valentin Schwind, Slim Abdennadher, and Albrecht Schmidt. 2018. I Like to Move It: Investigating the Effect of Head and Body Movement of Avatars in VR on User's Perception. In *Proc. CHI EA '18*. ACM, New York, NY, USA, Article LBW099, 6 pages. https://doi.org/10.1145/3170427.3188573

[43] Aqdas Malik, Amandeep Dhir, and Marko Nieminen. 2016. Uses and gratifications of digital photo sharing on Facebook. *Telematics and Informatics* 33, 1 (2016), 129–138.

[44] R. Mekuria, K. Blom, and P. Cesar. 2017. Design, Implementation, and Evaluation of a Point Cloud Codec for Tele-Immersive Video. *IEEE Transactions on Circuits and Systems for Video Technology* 27, 4 (April 2017), 828–842. https://doi.org/10.1109/TCSVT.2016.2543039

[45] Andrew D. Miller and W. Keith Edwards. 2007. Give and take: a study of consumer photo-sharing culture and practice. In *Proc. CHI '07*. Acm, 347–356.

[46] Niels Christian Nilsson, Stefania Serafin, and Rolf Nordahl. 2013. The Perceived Naturalness of Virtual Locomotion Methods Devoid of Explicit Leg Movements. In *Proceedings of Motion on Games (MIG '13)*. ACM, New York, NY, USA, 133:155–133:164. https://doi.org/10.1145/2522628.2522655

[47] Donald A. Norman and Stephen W. Draper. 1986. *User Centered System Design; New Perspectives on Human-Computer Interaction*. L. Erlbaum Associates Inc., Hillsdale, NJ, USA.

[48] Megan Norris and Luc Lecavalier. 2010. Evaluating the Use of Exploratory Factor Analysis in Developmental Disability Psychological Research. *Journal of Autism and Developmental Disorders* 40, 1 (01 Jan 2010), 8–20. https://doi.org/10.1007/s10803-009-0816-2

[49] Daisuke Okabe, Mizuko Ito, Jan Chipchase, and Aico Shimizu. 2006. The social uses of purikura: photographing, modding, archiving, and sharing. In *Pervasive image capture and sharing workshop, ubiquitous computing conference*. 2–5.

[50] D. Roth, J. L. Lugrin, D. Galakhov, A. Hofmann, G. Bente, M. E. Latoschik, and A. Fuhrmann. 2016. Avatar realism and social interaction quality in virtual reality. In *2016 IEEE Virtual Reality (VR)*. 277–278. https://doi.org/10.1109/VR.2016.7504761

[51] Maria Roussou. 2004. Learning by Doing and Learning Through Play: An Exploration of Interactivity in Virtual Environments for Children. *Comput. Entertain.* 2, 1 (Jan. 2004), 10–10. https://doi.org/10.1145/973801.973818

[52] Elizabeth B-N Sanders and Pieter Jan Stappers. 2012. *Convivial toolbox: Generative research for the front end of design*. BIS Amsterdam.

[53] Thomas Schubert, Frank Friedmann, and Holger Regenbrecht. 2001. The experience of presence: Factor analytic insights. *Presence: Teleoperators & Virtual Environments* 10, 3 (2001), 266–281.

[54] Valentin Schwind, Pascal Knierim, Cagri Tasci, Patrick Franczak, Nico Haas, and Niels Henze. 2017. "These Are Not My Hands!": Effect of Gender on the Perception of Avatar Hands in Virtual Reality. In *Proc. CHI '17 (CHI '17)*. ACM, New York, NY, USA, 1577–1582. https://doi.org/10.1145/3025453.3025602

[55] Pavica Sheldon and Katherine Bryant. 2016. Instagram: Motives for its use and relationship to narcissism and contextual age. *Computers in Human Behavior* 58 (2016), 89 – 97. https://doi.org/10.1016/j.chb.2015.12.059

[56] Mel Slater, Martin Usoh, and Anthony Steed. 1994. Depth of Presence in Virtual Environments. *Presence: Teleoperators and Virtual Environments* 3, 2 (Jan. 1994), 130–144. https://doi.org/10.1162/pres.1994.3.2.130

[57] Harrison Jesse Smith and Michael Neff. 2018. Communication Behavior in Embodied Virtual Reality. In *Proc. CHI '18*. ACM, New York, NY, USA, 289:1–289:12. https://doi.org/10.1145/3173574.3173863

[58] Marc Steen, Magnus Eriksson, Joke Kort, and Peter Ljungstrand. 2012. (PDF) D8.8 User Evaluations of TA2 Concepts. https://www.researchgate.net/publication/291351579_D88_User_Evaluations_of_TA2_Concepts

[59] Jonathan Steuer. 1992. Defining virtual reality: Dimensions determining telepresence. *Journal of communication* 42, 4 (1992), 73–93.

[60] Martin Usoh, Ernest Catena, Sima Arman, and Mel Slater. 2000. Using Presence Questionnaires in Reality. *Presence: Teleoper. Virtual Environ.* 9, 5 (Oct. 2000), 497–503. https://doi.org/10.1162/105474600566989

[61] D. T. van Bel, KCHJ Smolders, W. A. IJsselsteijn, and Y. A. W. de Kort. 2009. Social connectedness: concept and measurement. In *Proc. IE '09*. IOS Press.

[62] Nancy A. Van House. 2009. Collocated photo sharing, story-telling, and the performance of self. *International Journal of Human-Computer Studies* 67, 12 (2009), 1073–1086.

[63] Martijn Vastenburg, Natalia Romero Herrera, Daniel Van Bel, and Pieter Desmet. 2011. PMRI: development of a pictorial mood reporting instrument. In *Proc. CHI '11 EA*. ACM, 2155–2160.

[64] Arnold P. O. S. Vermeeren, Effie Lai-Chong Law, Virpi Roto, Marianna Obrist, Jettie Hoonhout, and Kaisa Väänänen-Vainio-Mattila. 2010. User Experience Evaluation Methods: Current State and Development Needs. In *Proc. NordiCHI '10*. ACM, New York, NY, USA, 521–530. https://doi.org/10.1145/1868914.1868973

[65] Froukje Sleeswijk Visser, Pieter Jan Stappers, Remko Van der Lugt, and Elizabeth BN Sanders. 2005. Contextmapping: experiences from practice. *CoDesign* 1, 2 (2005), 119–149.

[66] David Vronay, Shelly Farnham, and John Davis. 2001. PhotoStory: Preserving emotion in digital photo sharing. *Internal research paper* (2001).

[67] T. Waltemate, D. Gall, D. Roth, M. Botsch, and M. E. Latoschik. 2018. The Impact of Avatar Personalization and Immersion on Virtual Body Ownership, Presence, and Emotional Response. *IEEE Transactions on Visualization and Computer Graphics* 24, 4 (April 2018), 1643–1652. https://doi.org/10.1109/TVCG.2018.2794629

[68] Bob G. Witmer and Michael J. Singer. 1998. Measuring Presence in Virtual Environments: A Presence Questionnaire. *Presence: Teleoper. Virtual Environ.* 7, 3 (June 1998), 225–240. https://doi.org/10.1162/105474698565686

[69] Peter Zimmermann. 2008. Virtual reality aided design. A survey of the use of VR in automotive industry. In *Product Engineering*. Springer, 277–296.