Interactions across Displays and Space: A Study of Virtual Reality Streaming Practices on Twitch

LIWEI WU, University of Waterloo, Canada
QING LIU, University of Waterloo, Canada
JIAN ZHAO, University of Waterloo, Canada
EDWARD LANK, University of Waterloo, Canada

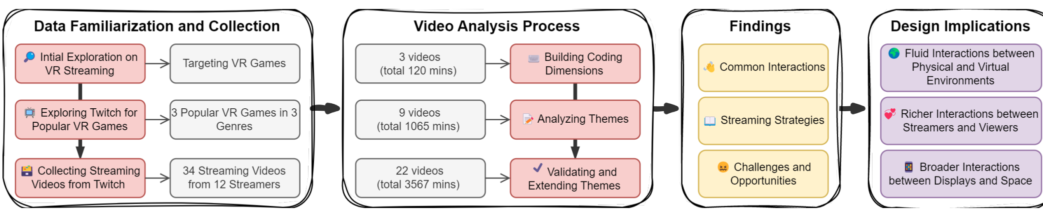

Fig. 1. The research process and key results of this study: we present a multi-phase thematic analysis of 34 streaming videos from 12 VR streamers on Twitch; we report common interactions, strategies, and challenges for VR streaming; and we present six design implications in three directions for enhancing VR streaming experience for both streamers and viewers.

The growing live streaming economy and virtual reality (VR) technologies have sparked interest in VR streaming among streamers and viewers. However, limited research has been conducted to understand this emerging streaming practice. To address this gap, we conducted an in-depth thematic analysis of 34 streaming videos from 12 VR streamers with varying levels of experience, to explore the current practices, interaction styles, and strategies, as well as to investigate the challenges and opportunities for VR streaming. Our findings indicate that VR streamers face challenges in building emotional connections and maintaining streaming flow due to technical problems, lack of fluid transitions between physical and virtual environments, and not intentionally designed game scenes. As a response, we propose six design implications to encourage collaboration between game designers and streaming app developers, facilitating fluid, rich, and broad interactions for an enhanced streaming experience. In addition, we discuss the use of streaming videos as user-generated data for research, highlighting the lessons learned and emphasizing the need for tools to support streaming video analysis. Our research sheds light on the unique aspects of VR streaming, which combines interactions across displays and space.

CCS Concepts: • **Human-centered computing** → **Empirical studies in HCI**; **Virtual reality**.

Additional Key Words and Phrases: live streaming, Twitch, video analysis

ACM Reference Format:
Liwei Wu, Qing Liu, Jian Zhao, and Edward Lank. 2023. Interactions across Displays and Space: A Study of Virtual Reality Streaming Practices on Twitch. *Proc. ACM Hum.-Comput. Interact.* 7, ISS, Article 437 (December 2023), 24 pages. https://doi.org/10.1145/3626473

Authors' addresses: Liwei Wu, a92wu@uwaterloo.ca, University of Waterloo, Waterloo, Canada; Qing Liu, qing.liu@uwaterloo.ca, University of Waterloo, Waterloo, Canada; Jian Zhao, jianzhao@uwaterloo.ca, University of Waterloo, Waterloo, Canada; Edward Lank, lank@uwaterloo.ca, University of Waterloo, Waterloo, Canada.

1 INTRODUCTION

Live streaming is increasingly popular. According to Twitch, there are more than 2.5 million viewerships at any given moment watching live streams on its platform, and a total of more than 1.3 trillion minutes were watched in 2021 [3]. Although live streams are often viewed on 2D displays, due to the proliferation of virtual reality (VR) devices, live performance and game experiences in digital spaces are also frequently live-streamed on streaming platforms [42]. For example, a streamer can play VR games wearing a head-mounted display (HMD) while viewers watch the stream on their 2D displays. With continuing investment from companies such as Meta [4], HTC, and Huawei in VR, the streaming of VR games and experiences will arguably become an increasingly integral part of the content library available on live streaming platforms.

Live streaming has been a popular research field for many years and there is significant past research focusing on different aspects of streaming including streaming interaction techniques [28, 35, 39, 66], streaming culture and community [19, 40, 41, 50, 63], and the social impact of streaming [22, 26, 37, 41]. However, few of these past studies focus on VR streaming. As a result, the literature is relatively sparse on aspects including the common interaction styles between streamers and their viewers, streaming strategies, and the challenges and opportunities for VR streaming.

Based upon an initial search of VR streaming on YouTube and Twitch, we find the streaming of VR games accounts for a majority of live VR streaming activities; all of the top-50 results from a search of VR live streaming on both platforms contain uniformly VR game streams. Thus, in this paper, we study streaming for VR games as a first step in the investigation of VR streaming.

VR streaming is unique as a combination of interactions both on the 2D displays and in the 3D space, and streaming for VR games is much different from streaming for traditional 2D/3D games. Throughout this text, we will refer to streaming for traditional 2D/3D games as "traditional streaming". For instance, there exist breakdowns and mismatches of environment setups between the streamers and the viewers. Specifically, the streamers are usually standing, wearing a VR head-mounted display (HMD), and holding controllers in their hands; whereas for traditional streaming, the streamers are usually sitting in front of normal displays and using a keyboard and mouse. In addition, the streamers in VR cannot directly control the streaming contents that the viewers see; whereas for traditional streaming, the streamers usually have full control of the streaming contents as they can easily create, change, composite, and delete the windows on the screen with mature streaming software (e.g., OBS). As a result, the streamers in VR need to map different streaming windows that are originally designed for 2D displays into VR to better interact with the viewers, e.g., windows for reading messages from the streaming chat, windows that represent real-world versus avatar movements, promotion information for revenue generation, etc. These differences make VR streaming unique and hold significant potential to inform the future design of interactions across displays and space, calling for further investigation and exploration in this field.

In this paper we present a multi-phase thematic analysis of 34 streaming videos from 12 VR streamers on Twitch (Table 1) [11]. The overall research process is shown in Figure 1. Our results distill a range of *interaction styles and strategies* that the streamers commonly use to interact with and engage their viewers including managing conversation tempos, building emotional connections with viewers, and the use of humor and jokes to capture and keep interest. We also identify the *challenges and frustrations* that the streamers face producing engaging VR streams, including difficulty switching between VR space and 2D displays, technical problems caused by VR hardware and software, and in-game scenes that could be better designed for streaming. *Design implications*, calling for future tools to better support streaming for VR, are explored based on our results. The contributions in this paper are two-folds:

1. Empirical knowledge of common interaction styles, strategies, and challenges for VR streaming.

2. Design implications for future design for more engaging VR streaming experience and better interaction techniques across displays and space.

2 RELATED WORK

Our work is related to the research on live streaming and virtual reality (VR), including the study of the current streaming practices and innovative interaction techniques to enhance the streaming experience. Also, our video analysis methodology is enlightened by past HCI studies using user-generated online videos as a source of study data.

2.1 Studies on Live Streaming

Online live streaming and events have been widely studied in the HCI communities recently due to the growth of the streaming industry. Live streaming and social events are widespread on various platforms such as Twitch [53], Facebook Live, Twitter Live, and Snapchat [27, 55]. Taylor conducted a comprehensive review on the transformation of broadcasting into live streaming as a form of media, exploring the associated opportunities and challenges [56]. Though gaming is the most popular content for streaming, the viewers also like to watch streaming contents that are educational [19, 26], inspiring [22], and unique [5]. It is shown that the immersion, immediacy, interaction, and sociality [27] provided by live streaming are the main reasons for the prosperity of live streaming, distinguishing it from the traditional media formats. VR, as an immersive media, naturally strengthens those attractive features of streaming. VR streaming raises our research interest because it shapes a unique combination of interactions between streamers and viewers involving both 2D displays and 3D space.

Many past studies have explored the motivations for viewers to watch live streaming. These motivations include social dynamics [24, 64, 65], culture and heritage [36], and monetary motivations [64, 65]. As Hamilton et al. pointed out, live streaming acts as a third virtual space where informal social communities can emerge and proliferate [28], which attracts many viewers for its socialization. In addition, most live streaming platforms provide meaningful methods for the viewers to interact with the streamers like live chat and Danmaku comments [38, 65, 72], as well as monetary rewards for the streamers [64]. These interactions between the viewers and streamers can help build social support for the viewers emotionally and instrumentally [64], but they are currently most happening through interactions on 2D displays, not fully taking advantage of the 3D space that VR provides.

Common strategies for streamers to attract and keep viewers in their streaming have also been widely studied. These studies include investigating the effects on the streaming experience from the influence of streamers' characteristics [22, 50], streaming contents [5, 19, 26, 37], and streamers' ways of expression [40]. As Pellicone et al. pointed out, appealing streaming is a combination of the uniqueness of the streamer's persona and the excellent integration of technological, social, and game-play-based skills [50]. Our research aims to investigate the unique aspects of VR streaming, which combines 2D displays and 3D space. We seek to understand the similarities and differences in streaming strategies compared to traditional streaming, leading to motivate our work.

2.2 Interaction Techniques for Streaming

Interaction is a significant aspect of live streaming to differentiate itself from online video watching. Past research on streaming interaction techniques demonstrated two broad directions: (i) providing better interaction context, and (ii) enhancing interactions' expressiveness and engagement.

In one streaming session, information is delivered through multiple channels including streaming text chat, voice conversations, and videos on the screen, which can be overwhelming for both streamers and viewers to follow and understand. By providing a better context for the streaming (e.g., easier-to-digest conversations and clearer relationships between streaming and conversation contents), communication quality between streamers and viewers can be elevated, which leads to a better streaming experience. Pan et al. has demonstrated a visualization tool for Twitch's chatroom to help streamers better understand the relationship between gameplay and the viewers' chatting behaviors [48]. Similarly, VisPoll [14] enables streamers to specify and aggregate the massive visual inputs from viewers to facilitate and enhance the interactions between streamers and viewers. Some other research demonstrated better recapitulation tools for the streaming contents for the viewers when watching learning contents in streaming both after [21] and during [39] the live streaming session.

The main interaction between streamers and viewers heavily relies on text-based conversations in the current streaming form [19, 28, 41], which certainly constrains the expressiveness of the viewers' emotions and their engagement with the streaming. To elevate that, past work explored enhancing streaming interactions with diverse multimedia. Glickman et al. provided an interface development toolkit to investigate the common challenges for audience participation in games during streaming [23]. Lessel et al. presented Helpstone [35], a live streaming tool for the card game *Hearthstone*, empowering the viewers to interact with the streamers through voting, sending alerts and card deck check on top of chatting. For the viewers with an intention to learn through educational streaming sessions, diverse multimedia including text, audio, video, image, stickers, and screenshots are enabled to support language learning [13] and collaborative social learning experience for online education [29]. However, these studies' exploration on multimedia may be constrained in the 2D contents presented on 2D displays. There is still much room to investigate how 3D content and 3D presentation methods like VR can create new streaming forms and enhance the streaming experience.

Though interaction techniques for streaming are widely explored in the past, all the streamers in the studies use 2D displays with mice and keyboards. There is still not enough understanding of the challenges and requirements for the interaction techniques between streamers and viewers when the streamers are using VR HMD and VR controllers, which intrigues us to further investigate in our study.

2.3 Streaming in Virtual Reality

Since affordable high-performance VR devices and well-crafted VR games have only begun to be in the consumer market in recent years, the research on VR games, especially VR game streaming, though beginning to catch increasing interests [43, 68, 70], is relatively sparse.

Several studies have investigated interaction techniques for streaming in VR, but their primary focus is not on supporting streamers to produce engaging streaming sessions for massive viewers. For example, education, by nature, is benefited from immersive multi-medium and is widely explored using VR and streaming technologies to enhance the learning experience for the learners [46, 57, 58]. Simulation and training is another field where streaming and VR technologies are used together to facilitate and enhance the training process in multiple scenarios [54, 62, 67]. For all these cases, streaming is mainly used as a visual communication method between the users in VR and other users without VR. Particularly, numerous recent studies have investigated "mixed-screen" collaborations where users external to the VR environment engage with users who are immersed in a VR headset [7, 17, 25, 49, 59]. This setup emulates a shared streaming experience, as non-VR users can observe and interact with their VR counterparts, but it remains confined to a one-to-one or one-to-few paradigm and can not be directly used for VR streaming as a mass medium. Additionally, several recent studies have also explored enhancing the VR streaming experience by making streaming content more engaging for viewers [31, 60]. However, few studies have focused on the common strategies and interaction techniques for VR streaming as an online live media broadcasting activity for massive viewers to watch.

In another aspect, observation on the user behaviors and the shared visual contents during streaming has a vital role in streaming related studies since it provides important context information to help the researchers better understand the challenges and frustrations the users come across [5, 33, 51]. Due to similar reasons, for social VR related studies, observation helps the researchers build connections between the user behaviors in the physical environments and the visual contents in the virtual environments displayed in the HMD, which is also widely adopted in many recent studies in HCI communities. These observations include indirect observations on the user-generated contents to understand user behaviors through posts in online forums [69] and posted online videos [15], as well as direct observations on the user behaviors in live VR activities [52] or designed experiment environments [44]. From the streaming viewers' perspective, recent studies have also noticed the different influences on the viewing experience between first-person and third-person observation perspectives for the VR games being played during streaming [18]. Enlightened by these past studies, our study managed to investigate VR streaming through observations on recorded streaming videos and conversations in the streaming chat panels.

2.4 User-generated Online Videos as Sources of Study Data

User-generated online videos, especially YouTube videos, are getting increasing attention from the research community as an important source of data to observe and analyze users' behaviors. This is rooted in Anthony et al.'s work on using YouTube videos paired with surveys to research the use of touchscreens by people with motor impairments [6]. We see more research utilizing user-generated videos in recent years to understand the user behaviors in different research topics including: insertable device interactions [34], use of tablets by young children [32], and interactions when cooking in the kitchen [47]. Particularly, Dao et al. analyzed 223 YouTube videos to investigate VR failure outside of lab settings and demonstrated the potential for analyzing user behaviors in VR with videos [15]. Differently, our study investigates the general process and common strategies for the streamers in their professional streaming practices, rather than focusing on one particular scenario (i.e., VR failures in daily usage).

The past research has enlightened our study to use user-generated streaming videos on Twitch along with the chat texts during the streaming to investigate and understand the VR streamers' behaviors. In addition, since these streaming videos are usually generated by directly recording the streamers' entire sessions without further editing, they could serve as a more objective and better quality data source compared to YouTube videos for recording and capturing streamers' behaviors, as well as their details of interactions with the viewers.

3 METHOD

As an integral aspect of the streaming session, interactions between streamers and their streaming devices, as well as between viewers and their viewing devices, inevitably impact the overall streaming experience. This study primarily focuses on the human-human interactions specifically between VR streamers and their viewers, building upon the aforementioned human-device interactions. In this research, we aim to enhance our understanding of the current interactions, streaming strategies and challenges of VR streaming by exploring the following three research questions:

RQ1 - Interactions: a) What are the common interactions between the streamers and the viewers? b) How about other streaming participants (e.g., other game players)?

RQ2 - Strategies: What are the common strategies the streamers use to engage their viewers and improve their watching experience?

RQ3 - Challenges: What are the frustrations and challenges that the streamers have in their current streaming workflows?

To investigate VR streaming practices, we initially conducted observations using live streaming videos. While this method is commonly used for non-streaming studies, it has also found application in streaming-related research [16, 71]. This approach provides a comprehensive and objective view of VR streaming, allowing us to examine game content, streamer behavior in the physical space, and viewer reactions simultaneously. Compared to surveying or interviewing streamers and viewers, studying streaming videos alongside chat texts offers advantages in understanding the integrated and cohesive flow of VR streaming. While we watched the streaming videos from a third-person perspective, the content viewed closely resembled what streamers saw through their head-mounted displays (HMDs). Additionally, streamers expressed their thoughts aloud during streaming, similar to a "thinking aloud" process. Thus, by analyzing video content, streamers' dialogues, and text messages in the streaming channel, we can collect ample data to address our research questions. Our overall data exploration approach was based on Braun and Clarke's [11] description of thematic analysis. We started by exploring VR streams on popular platforms and watching multiple live-stream videos to familiarize ourselves with the subject. Next, we coded three 40-minute video clips from three of the 12 streamers to generate initial codes. Utilizing these initial codes, we performed inductive coding on videos from the remaining nine streamers to identify overarching themes in our data. Finally, we extended our analysis with 22 additional videos including both VR streaming and traditional streaming content from the 12 streamers to validate our initial thematic analysis and compare the differences between VR streaming and traditional streaming.

3.1 Data Familiarization and Collection

To develop an initial understanding of VR streaming, we first conducted searches on YouTube with the keywords "VR" and "streaming", and we found that all the top-50 results were about streaming VR games or reviews of VR equipment. We then conducted additional searches with the term "VR" on Twitch, one of the largest live streaming platforms, and also found that all of the top-50 results were about streaming VR games. Thus, as an initial step in studying practices and challenges of VR streaming, we focus primarily on VR game streaming videos in our study data as they appear to represent a significant fraction of VR streaming content.

After watching a number of live-stream videos to develop an appreciation for the medium, we collected our data set for analysis. We collected 34 streaming videos (4,752 minutes in total length) from 12 VR streamers hosted on Twitch. To search for streaming videos representative of VR streaming practices, we focused on the three most popular VR games during our data collection period on Twitch. These games span various genres and showcase distinct VR features (Figure 2): *Beat Saber*, a music rhythm game requiring vigorous body movements; *Half-life Alyx*, a narrative-driven first-person shooter game with immersive interactivity; and *VRChat*, a social VR game recognized for its extensive avatar and environment customization options, as well as multiplayer engagements. To better portrait the diversity of VR streamers and capture the wide spectrum of VR streaming practices, we purposefully curated streaming videos from streamers with various represented genders, levels of popularity (measured by follower counts), and streaming setups. We first chose the top two recommended videos and their corresponding streamers when searching the game titles on Twitch. Additionally, to enrich the diversity of streamers' characteristics and streaming styles, we manually pick up two additional videos and their streamers per game based on their search result rankings. Table 1 indicates the profiles of the 12 streamers in our collection. Among the 12 streamers, five (P3–6, P11) used a physical camera, five used self-costumed 3D avatars (P2, P9–12), two used 2D avatars (P1, P8), and one streamed without showing himself (P7).

3.2 Video Analysis Process

Streaming videos are relatively long (usually 2 to 4 hours). To speed up coding, we first captured the videos and uploaded their audio tracks to Otter.ai [1] to extract transcripts with time tags. We then applied inductive coding using the transcripts supplemented with video data. Specifically, we began by reading transcripts, and when we came across interesting moments or found concepts ambiguous in the transcripts, we referred back to the videos with the time tags and watched short periods of the videos to code these moments. Overall, the analysis of the VR videos in our study proceeded in three phases: (i) coding dimension generation, (ii) inductive coding, and (iii) validation and extension.

As the first step, we analyzed one 40-minute video clip for each game to generate coding dimensions. These three video clips were first analyzed by the primary researcher to generate the initial coding dimensions. Then, a second member of the research team watched, discussed, and revised the dimensions until the team reached a consensus. Finally, the overall research team discussed the overall coding dimensions, focusing on effectively addressing the three research questions proposed earlier.

Once the coding dimensions had been set, the initial videos were fully coded using the final coding dimensions and the results were grouped into four themes (Table 2): (a) meta info: general information about this streaming video, (b) gaming interactions: how the streamers interact with the games, (c) non-gaming interactions: how the streamers interact with other parts including the

Fig. 2. The three selected popular VR games on Twitch: (a) *Beat Saber*, a music rhythm game with 619K followers; (b) *Half-Life Alyx*, a first person shooter (FPS) game with 71.4K followers; and (c) *VRChat*, a social game with 2.9M followers. Images are from roadtovr.com

Table 1. The 12 streamers and their videos involved in our study. Column 2 shows the 12 fully-coded videos from the first and second phases of our analysis, in which ⋆ indicates the ones analyzed in the first phase. Column 3 shows the 22 extra videos analyzed in the third phase, in which the number in the parenthesis indicates the number of videos analyzed. See Section 3.2 for details.

ID	# of Followers	Fully-Coded Video	Extra Videos
P1	1.8K	Beat Saber, 40 min ⋆	Minecraft (2), 139 min; Craft Lesson (1), 208 min
P2	27.1K	Beat Saber, 82 min	Bug Fables (1), 315 min; Portal2 (1), 133 min; Minecraft (1), 110 min
P3	1.3K	Beat Saber, 40 min	GTA5 (1), 19 min; Minecraft (1), 74 min
P4	55.4K	Beat Saber, 165 min	Beat Saber(4), 524 min
P5	64.4K	Half-life Alyx, 40 min ⋆	N/A
P6	9.8K	Half-life Alyx, 130 min	Half-life Alyx (2), 252 min; Halo Infinite (1), 287 min
P7	603	Half-life Alyx, 110 min	N/A
P8	13	Half-life Alyx, 40 min	N/A
P9	263k	VRChat, 40 min ⋆	VRChat (2), 372 min; Chatting (1), 159 min
P10	94.1k	VRChat, 196 min	VRChat (3), 746 min; Chatting (1), 229 min
P11	77.9k	VRChat, 149 min	N/A
P12	1.3k	VRChat, 153 min	N/A

Table 2. The final 18 dimensions for coding the streaming videos, grouped in four themes.

Meta Info	
Collaboration Type	*how do the streamer(s) produce the streaming and collaborate?*
Main Streamer Type	*what are the characteristics for the main streamers?*
Play Style	*how do the streamers play the game?*
Conversation Style	*how do the streamers communicate with the viewers?*
Streaming Composition	*how are the streaming components mapped on the screen?*
Gaming Interactions	
Game Content	*what gaming elements do the streamers talk/discuss about?*
Game Behavior	*how do the streamers explain their movements and strategies?*
Game Performance	*how do the streamers talk about their performances while playing?*
Game Experience	*how do the streamers express their like/dislike about the game?*
Emotion Expression	*how do the streamers express their emotions while playing?*
Non-gaming Interactions	
Streaming Management	*what do the streamers do to maintain the streaming functions?*
Conversation Topics	*what non-gaming topics do the streamers usually talk about with the viewers?*
Interaction with the Viewers	*what are the common interactions between the streamers and the viewers?*
Strategies to Engage the Viewers	*what efforts do the streamers try to engage the viewers in the streaming?*
Interactions with Non-viewers	*What are the common non-viewer roles and their interaction patterns?*
Frustrations / Opportunities	
Raw Usages of the Equipment	*what are the cases the streamers use the equipment in an unexpected way?*
Bad Experience	*what are the situations where the streamers or the viewers may feel bad?*
Undesigned Scenes	*what are the common scenes in VR that can be designed to be more engaging?*

viewers and other players, (d) frustrations and opportunities: how current tools fail the expectations of the streamers and what can be improved for better streaming experience.

The primary researcher then coded additional nine complete streaming videos (length varies from 1 hour to 3 hours), three for each game (1,065 minutes in total length), based on the coding dimensions generated above. Themes were refined throughout this coding step. A second researcher in the research team then watched the twelve fully-coded videos, checked the codes, and discussed disagreements with the primary researcher to modify the codes until full consensus. This resulted in the 12 fully-coded videos in Table 1.

In the final phase, two of the researchers in the team each watched two to four more streaming videos from seven of the 12 streamers whose videos were initially coded to further confirm the codes and thematic analysis. As shown in Table 1, this added an additional 22 videos (of around 3,567 minutes in total length) to our data set. To better understand the streaming process and distinguish VR streaming, we analyzed a diverse range of videos. These included VR games (initially selected and others), 2D/3D games, and craft lessons. This allowed us to make meaningful comparisons between VR and traditional streaming, facilitating a nuanced exploration of their unique characteristics.

4 FINDINGS

In this section, we report our findings structured around our research questions (Figure 3). We start by reporting common streaming interactions among the streamers, the viewers, and other streaming participants. Then we report the common strategies the streamers use for engaging their viewers during VR streaming. Finally, we highlight the observed challenges and frustrations for the streamers during their VR streaming sessions.

4.1 RQ1-a: Common Interactions with Viewers

Currently, Twitch enables interaction between viewers and streamers primarily through streaming chat, resulting in text or voice-based conversations as the predominant form of interaction during streaming sessions. These conversations can be categorized into two groups based on their purpose: (i) sharing the game experience with the viewers (i.e., *game related interactions*) and (ii) hosting the streaming sessions for the viewers (i.e., *session hosting interactions*).

4.1.1 Game Related Interactions. For game related interactions, there are five common conversation topics: (1) *game contents on the screen*, (2) *player status in the game*, (3) *movement plans and explanations*, (4) *self-evaluation*, and (5)*comments and critiques*.

Streamers in VR games, like those in 2D/3D games, describe in-game elements such as objects, environments, and characters while playing. Streamers typically describe their current in-game status and explain how it influences their gameplay. They focus on providing reasons for past movements and future plans, such as saying, *"I'm going to have some fun with you. And then shoot this glass and a grenade in a room in a minute."* -P6. Streamers' descriptions are crucial for viewers to comprehend on-screen streaming content, especially when viewers may not have complete visibility of these descriptions on their screen.

Moreover, the streamers sometimes share the self-evaluation of their in-game playing performance. This usually happens when the streamers do extremely well or poorly in the games and is communicated via strong emotional expressions, e.g., *"I'm not that good at Beat Saber but I do have a lot of fun playing it so I'm really I feel like a champion even though I know that I'm really not that good being said."* -P2. Finally, the streamers also share their comments and critiques about the game or the game related content with the viewers. This usually includes what they like and dislike about the games being played and how they would like to see them change. For example, when P2 was playing a badly designed level in Beat Saber, he commented *"I'm so confused. What is this? So much mapping [on how the level is designed in accordance with the beats in the music] is really good and something is really wack."*

4.1.2 Session Hosting Interactions. For session hosting interactions, streamers actively manage the stream chat and frequently chat with the viewers. The conversations' contents can be categorized into five groups: (1) *controlling streaming flow*, (2) *appreciating the viewers*, (3) *making requests*, (4) *referencing third parties*, and (5) *expressing streaming problems*.

In VR streaming, controlling the flow of the stream is a common non-gaming interaction. This includes announcing the start, end, and breaks of the streaming sessions, as well as recapping and

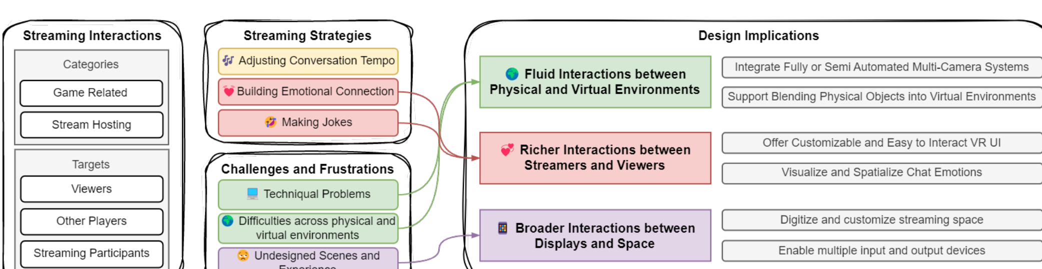

Fig. 3. In accordance to the research questions, our research findings include (a) *Streaming Interactions*, (b) *Streaming Strategies*, and (c) *Challenges and Frustrations*. Based on the findings, we introduce six design implications in three directions: (a) *Fluid Interactions between Physical and Virtual Environments*, (b) *Richer Interactions between Streamers and Viewers*, and (c) *Broader Interactions between Displays and Space.*

sharing plans for future sessions. VR games can be physically intense and the streamers need to take frequent breaks for rest, providing them with more opportunities to recap and plan during their streams. Periodic recaps during VR streams summarize the events of the current session, helping new viewers understand the context, whether they are newly joining viewers or new to the channel. Additionally, sharing plans informs viewers about upcoming events both within the current session and in future sessions. For example, P12 said *"We're just chilling, you know, hanging out with our friends usually, playing murder, playing lots of lots of games"* at the beginning of the streaming session to share a brief plan with the viewers and said *"We get a little tired playing murder so we're gonna move on to our next activity and that is chilling in our rooftop."* to both recap what he did and move to the next activity.

Another important type of non-gaming interaction is appreciating the viewers, which is common in streaming. The streamers express their appreciation to the viewers who join their streaming sessions, follow their streaming channels, and make donations or send gifts. If the streamers want to appreciate many viewers at the same time, they may express their appreciation as a group like *"Thank you so much, Mr.Midnight, Stargate and Gigi Nord for the followers, guys, welcome to the children's (the players') live streaming."* -P2 but most streamers still prefer to appreciate the viewers individually, e.g., *"Amantha thank you. I always call out. I feel you are fucking nice to me today. Hell yeah. Thank you."* -P7.

To maintain an engaging streaming chat, streamers occasionally ask viewers to adhere to certain guidelines. These requests may involve refraining from discussing specific topics or focusing on the content being streamed. For example, when a viewer shared an offensive joke, P7 asked everyone to stop discussing it, stating, *"I won't entertain any comments about breast milk from now on."* Streamers may also request favors from viewers, such as providing feedback, advice, or visiting certain websites, which echoes the branding strategy observed by Pellicone and Ahn [50]. Streamers also discuss their streaming setups, including GPUs, VR headsets, and streaming software that supports VR, more prominently than traditional streaming. They may mention third-party resources, such as specific brands or products, when talking about their equipment or recommending tools to viewers. In cases of emergent problems during streaming sessions, streamers promptly inform viewers. For example, when P4 encountered a tracking issue, she immediately shared her concern with viewers, saying, *"What's wrong with my game? I feel weird. I'll try adjusting my position here."*

As presented above, a successful VR streaming session requires the streamers constantly switch contexts to handle various tasks. Compared to traditional streaming, the context switch for VR streaming is more difficult, not only jumping among UIs of different software on the screens but also across various devices between physical and virtual environments.

4.2 RQ1-b: Common Interactions with Other Streaming Participants

Streamers primarily interact with the streaming viewers, but they also engage with *other players* and *streaming co-hosts*.

4.2.1 Interactions with Other Players. If the game is a multiplayer game, the streamers inevitably interact with other game players. In our study, we observed that the streamers interact with other players in VRChat mostly by making fun movements and building interesting conversations. These players can be categorized into three groups: (1) *guests* who are invited by the streamers to play the game together; (2) *players informed* by the streamers that they are joining streaming, and (3) *players fully unaware* of being captured in streaming. Streamers may also invite their streaming viewers to join them as guests like *"By the way, guys, if you do have VRChat, and would like to play with us, you are more than welcome to come in."* -P12.

For the guests and informed players, they collaborate with the streamers actively to make the streaming interesting for the viewers to watch. Most of the time, these collaborations are in the form of a group conversation or an in-game small activity and the contents of these conversations and activities are usually pre-planned by the streamers. Many uninformed players treat streamers as regular players while playing the game, but some are confused by the streamers' actions, particularly when they interact with viewers through chat with noticeable body movements. We also observed one uninformed player becoming uncomfortable and even angry when discovering they were being streamed without prior notification from the streamer.

4.2.2 Interactions with Co-Hosts. Another group of non-viewers whom the streamers actively interact with is the co-host or moderator for the streaming session. In one of the fully coded streaming videos (P5), there was a female co-host. Similar to traditional streaming moderators [22, 38, 39, 50], her primary responsibility was to enhance communication by reading out interesting messages and questions in the chat to the streamer. This allowed the streamer to concentrate on the game without constantly removing their VR headset to check and respond to the chat.

For example, P5 asked the co-host *"Is there anybody else talking in the discord? I actually can't see the discord at all when I'm in VR,"* and the co-host checked the discord server and replied *"Nope, no one else,"* to help the streamer maintain the chat channels. In addition, she answered simple questions posted in the chat, such as what is the game being played, what is the streamer's discord server, and the time schedule for this streaming. As a partner, the streamer also built interesting conversations with her to make the streaming session more active and engaging for the viewers.

Interestingly, the "co-host" may be pets. In another fully coded streaming video (P10), the streamer placed a cat in front of the camera and played with the cat from time to time while she had a break during the game and when she finished the streaming session. We observed that the viewers were highly interested in the cat and what the cat was doing, which prompted a lot of conversations and interactions between the streamer and the viewers, e.g., *"Cheeto (the name of the cat), come back. Cheeto, come back right now! Bitch right that's it, that's it. How does he leave us? How does he leave us like that? (The cat got out of the camera and played by himself)".*

In general, interacting with non-viewers usually highly entertains the viewers, especially when unexpected moments happen. Thus, streamers usually welcome other participants in their streaming and must use various strategies to ensure their streaming sessions remain interesting and engaging.

4.3 RQ2: Common Streaming Styles and Strategies

As mentioned earlier, interaction between streamers and viewers heavily relies on texts [38, 72], given the current limitations of streaming platforms. In addition to the text-focused strategies such as moderator involvement, supportive viewers, and chatbot assistance for content monitoring [50], our study identified various conversation styles and voice-focused streaming strategies employed by VR streamers.

4.3.1 Conversation Styles. Three main styles for streamers to make conversations with the viewers were observed: (1) *monologue*, (2) *one-to-one*, and (3) *broadcast*. In the monologue style, the streamers talk like they are talking with themselves or they are talking towards no specific targets. This usually happens when the streaming chat is not active or the streamers are focused on game play.

One-to-one is a conversation style where the streamers are directly talking to one specific viewer or several viewers. These conversations typically begin with questions posed by viewers in the streaming chat, and streamers respond to these questions once they see them in the chat panels in VR. Some streamers may rephrase the viewers' questions to provide context for other viewers. One-to-one conversations are generally brief, with streamers offering concise responses consisting of one or a few sentences. These conversations rarely develop into longer exchanges with subsequent questions and responses. This limited interaction may be due to challenges in reading unoptimized texts in VR and the streamers' inability to sustain prolonged focus on the text panel while being immersed in VR.

Finally, broadcast, refers to the style that the streamers talk targeting every current viewer. This type of conversation usually starts with some reminding words like "Watch" or "Hey, chat!" to call for the attention of the viewers and consistently use second person phrases like "you" and "your" in the conversation to make the viewers feel the streamers are talking to them.

4.3.2 Streaming Strategies. Based on the three conversation styles, we further identified three main strategies widely used by VR streamers to better engage their viewers: (1) *adjusting conversation tempos*, (2) *building emotional connections* and (3) *making jokes*. Keeping the same conversation style and game contents for a long session can make the viewers feel bored. As a result, the streamers switch their conversation styles in accordance with the gaming contents to keep the viewers engaged or to re-engage the viewers. Streamers often use a monologue or broadcast style while playing games to keep viewers engaged, especially when they can't frequently check the chat panels in VR. During breaks or when the game is loading, they switch to a one-to-one style to interact with viewers before returning to the game or switching to another one. This long-short-long pattern helps maintain a predictable tempo for viewers to follow. It is also important to allow streamers to take short breaks, as VR games can be more physically demanding compared to 2D/3D games.

Emotional connection has been identified as a crucial factor for viewers in streaming [40, 63, 64]. However, establishing an emotional connection in VR streaming can be challenging when streamers' facial expressions are obscured by VR headsets, making the viewers hard to precept the streamer's emotions directly. Despite this obstacle, streamers strive to create emotional bonds by leveraging their virtual avatars and engaging in conversations to foster closeness with viewers. Consequently, streamers must rely on strong and exaggerated expressions in order to captivate viewers and evoke their emotions, which is also complemented by exaggerated body movements and frequent expression changes of their virtual avatars. For example, while P1 played one relatively difficult level in Beat Saber, he yelled *"You guys can see that you are here. When I passed this level what are you guys gonna say? Where are you at?"* to the viewers to also make them feel excited.

As a specific way of exaggerated expressions, joking is a prevalent and effective method used by streamers to entertain, engage, and evoke emotions in viewers. Some jokes describe the game contents in a funny way. Additionally, streamers use two significant types of jokes: teasing viewers with seemingly offensive remarks or challenging questions, and playfully accusing or blaming themselves. These types of jokes serve to reduce the perceived distance between streamers and viewers, fostering a closer connection between them. For example, P10 had a conversation with her viewers about how to eat a dish in Michelin-starred restaurants and some viewers said they enjoy a dish by eating each ingredient individually. P10 commented *"Okay, I'm sorry, you fuck up the dish that way. Okay, you're supposed to combine them because otherwise they suck. The dish fucking sucks if you do them individually, okay?"*

As a result of applying these three strategies, the streamers transform the streaming sessions from a serious performance into a casual, mutual conversation. VR streamers, in particular, require enhanced skills to effectively manage the dynamic between intense VR gameplay and engaging breaks, while also fostering emotional connections with viewers given the physical limitations from the VR headsets.

4.4 RQ3: Challenges and Frustrations for VR Streaming

We observed different failures and frustrations that streamers experienced based on the streamers' dialogues and interactions during their streaming sessions. These include both technical and viewing experience problems.

4.4.1 Technical Problems. All the streamers suffered from technical problems during their streaming sessions. The most frequent types of problems we observed included: (1) *VR tracking system problems*, (2) *low refresh rate*, and (3) *software failure*.

Many streamers experienced VR tracking problems and some of them suffered from this problem multiple times during a session. For example, the streamers suddenly find the VR controllers stop working or the virtual hands do not match the controllers. Another problem related to the hardware is the sudden drop of the refresh rate inside the VR HMD. The streamers usually notice this problem immediately because a low refresh rate can easily make the players feel uncomfortable and disoriented. Finally, many streamers use third-party software and tools to facilitate streaming for VR. Typical functions of these tools are mirroring desktop screens in VR and costuming in-game avatars, objects, and environments. However, some functions supported by third-party software may stop working or break the streaming session. For example, Beat Saber allows the players to make and import customized sabers and environments which may conflict with each other when used together and slow down or even crash the game.

If the problems are caused by the hardware or the streaming software, the streamers usually need to take off the VR HMD to work with the computers. If the problems are caused by the games, the streamers usually try different in-game configurations in the games' setting panels. Because streaming is a live, real-time activity, technical problems disrupt the streaming session significantly.

4.4.2 Viewing Experience Problems. Beyond technical problems, there is one major frustration we observed for nearly every streamer: *an inability to show the objects in the physical environment and the virtual environment at the same time* (Figure 4 a). This frustration pushes the streamers to frequently explain themselves to the viewers, in order to provide context and help viewers understand their actions. For example, in VR, the streamers cannot show how they interact with their own bodies or physical objects such as touching their legs or grabbing a water bottle. The streamers also cannot use their fingers or physical pointer to point at virtual objects in VR for the viewers. In both situations, the streamers have to explain to the viewers that they cannot see something the streamers are interacting with or they cannot see the virtual objects the streamers are talking about, which impacts the experience for both the streamers and the viewers.

Another important observation is that *there exist many undesigned scenes and experience* (Figure 4 b). Similar to many 2D/3D games [53], most of the VR games are not designed with considerations for streaming. We use "undesigned scenes" to refer to the VR game scenes that have the potential to be better designed for streaming purposes, which could improve both the streamers' and the viewers' engagement and experience. We categorized three groups of common undesigned scenes: (1) loading and saving screens, (2) scenes with camera angles too close or too far to the gaming objects, and (3) blank scenes when the streamers are not in VR. Depending on the games and the hardware that the streamers use, the loading and saving scenes can take several seconds to several minutes to finish. Currently, when loading and saving, there are only static images and the streamers usually try to talk with the viewers to make the waiting process less boring. Another typical undesigned scene is the camera angles in VR which are too close or too far for the viewers to understand the context and the game contents. These camera angles are designed for the players in VR but may not be comfortable and reasonable for the viewers to watch the game. Finally, sometimes the streamers need to step out of the stream and leave the headset and controllers on the

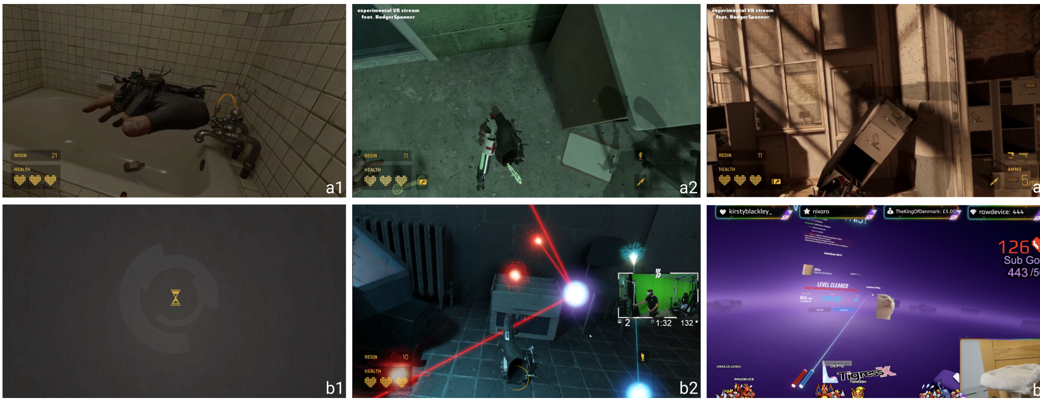

Fig. 4. Examples of frustrations caused by (a) being unable to show objects in physical and virtual environments at the same time as well as (b) undesigned scenes. (a1) The streamer is opening a bottle of water; (a2) The streamer is injecting the virtual object into his leg; (a3) The streamer is pointing at left with his finger; (b1) The loading screen for Half-life Alyx; (b2) The streamer is solving a puzzle but the camera is too close for the viewers to see the context; (b3) The streamer is absent from the scene, chatting with the viewers in front of the PC.

ground. In this situation, the viewers can only see a blank scene with a strange camera perspective caused by the loss of control from the streamers.

5 DISCUSSION

In this section, we begin by presenting the design implications derived from our findings. Next, we share our learned lessons about analyzing streaming videos. We then discuss the limitations of our study and suggest potential directions for future research.

5.1 Design Implications

In our results, we note technical challenges faced by streamers. Many of these challenges—poor tracking, poor in-VR views—are obvious targets for enhancement, and work on any of these represents obvious avenues to improve the VR streaming experience. For a smooth streaming experience, game and streaming software integration is essential. This is especially critical for VR streaming, where streamers need seamless control over both game content and streaming flow. Currently, these aspects are handled separately by game designers and streaming application developers, resulting in a lack of deep integration. Anticipating this need, we expect increasing collaboration between these two parties in the future. In this section, we provide six design implications for VR game designers and streaming application developers based on our findings (Figure 3) in three directions (Figure 5) to create more integrated and engaging VR streaming experience for both streamers and viewers.

5.1.1 ***Enable Fluid Interactions between Physical and Virtual Environments.*** In Section 4.4, we discussed one unique challenge of VR streaming, where streamers have to constantly switch between physical and virtual environments for breaks and bug checks. Moreover, they are tasked with presenting viewers with both the physical aspects of interaction, such as controller usage and movement, and the corresponding actions within the virtual environment. This disconnection between the two environments often forces streamers to alternate between them or provide detailed explanations for one environment while viewers focus on the other. To address this challenge, an ideal solution would be a system that automatically maintains and shares the streaming contexts in both the physical and virtual environments. This would alleviate the need for frequent switching and give rise to the following two design implications.

Integrate Fully- or Semi-Automated Multi-Camera Systems (A1). The utilization of multi-camera systems has been proven effective for VR content creation [45, 46]. Currently, most VR streamers manually control one physical camera to capture their physical environments and the default virtual camera in game for streaming contents. Introducing fully or semi-automated multi-camera systems that incorporate both physical and virtual cameras can enable streamers to seamlessly control multiple cameras simultaneously and facilitate sophisticated composition of views. This system will also enable personalized viewing experiences for the viewers tailored to individual preferences and comfort [18]. In addition, the multi-camera systems can auto-switch different cameras based on preset scenarios (e.g., when streamers take off headsets, switch to the main physical camera) or triggered by easy inputs from the streamers (e.g., use voice commands or gestures to switch to different cameras). Functions like auto tracking, autofocusing, and dynamic exposure can also be added to the system to further facilitate the automation of controlling cameras. In this way, the streamers can fluently and effortlessly control and compose multiple physical and virtual cameras at the same time, which allows the streamers to concentrate more on the streaming content, thus providing a better experience for the viewers.

Support blending physical objects into virtual environments (A2). Blending virtual and physical environments has been explored in VR research [46, 57, 73], but it is not currently supported in existing VR games or streaming software. However, features like pass-through, available on devices such as the Oculus Quest [61], already allow users to see their physical surroundings in

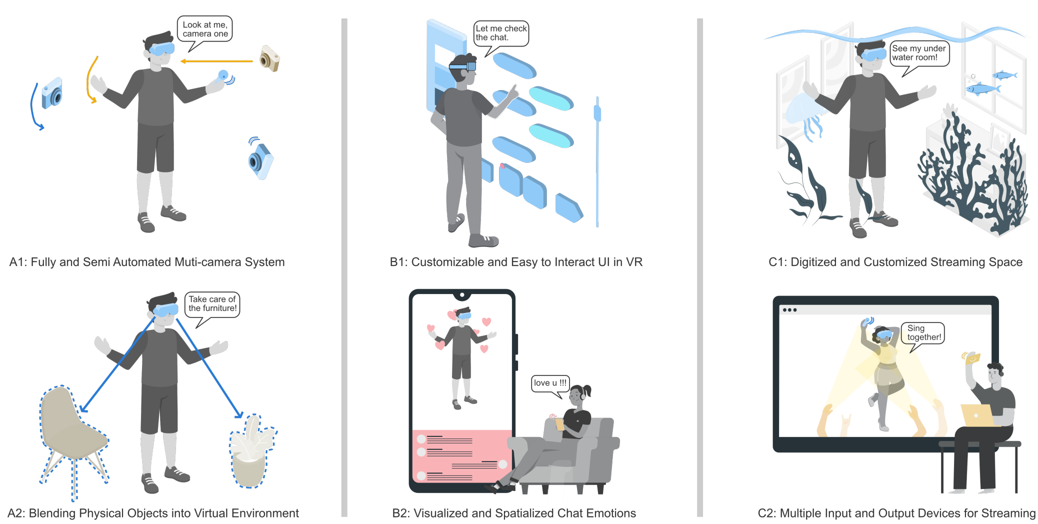

Fig. 5. Design implications for future VR streaming tools: (A1) Integrate fully- and semi-automated multi-camera system including physical and virtual cameras that can be easily controlled by streamers through voice command, hand gestures, and VR controllers; (A2) Support blending physical objects into virtual environments to allow the viewers to better get streaming contexts; (B1) Offer customizable and easy to interact UI specifically optimized for VR experience; (B2) Visualize and spatialize detected motions in streaming chat to build stronger emotional connections; (C1) Empower streamers to digitize and customize their streaming space for better self-expression and personal branding; (C2) Enable multiple devices as input and output devices for a more engaging experience.

real-time while experiencing virutal content. On the other hand, by utilizing improved cropping and masking techniques, a blending system could be developed to incorporate physical objects captured by physical cameras into VR scenes. This system can also prioritize streamers' privacy by only capturing specific objects (e.g., pets or water bottles) or activating when predetermined events occur (e.g., when streamers are not using controllers and their hands are touching physical objects). Future tools can work on the blending system, enabling streamers to seamlessly integrate physical objects into the virtual environment. This can eliminate the need for explicit explanations of transitions between virtual and physical environments and provide viewers with a more seamless viewing experience.

5.1.2 ***Facilitate Richer Interactions between Streamers and Viewers.*** As discussed in Section 4.3 and underscored by prior studies [24, 64, 65], establishing social dynamics and emotional connections between streamers and viewers is crucial for the engagement and success of streaming sessions. However, comparing to traditional streaming, VR streamers face challenges in building emotional connections with viewers due to heavy reliance on text-based communication, which is not well supported in VR and difficult for streamers to read, as well as the physical limitations imposed by VR headsets. To address this issue, future tools should enhance existing text-based communications and introduce additional features that foster stronger emotional connections.

Offer customizable and easy-to-interact VR UI (B1). Currently, VR streamers commonly rely on LIV.tv's customizable in-headset stream chat widget to access text conversations and receive alerts through HUD popups [2]. However, this chat widget simply maps the 2D chat UIs from a regular display, leading to two primary issues: the text size can be too small for streamers to read comfortably, and reading the texts may require body movements that appear strange to other players without proper context. Consequently, it can be more difficult for the streamers to juggle responsibilities like overseeing the chat, interacting with viewers, and concurrently participating in the game, which is a challenge already formidable in 2D/3D gaming scenarios for streamers [20]. To address these challenges, a more VR-adapted and flexible chat UI is required. This UI should enable streamers to easily view and interact with the streaming chat. For instance, UI elements such as text labels, buttons, and sliders should be larger in size to enhance readability and provide better interaction affordance in VR. In addition, these UI elements should support pointer-based interactions (e.g., laser or ray cursor), controller-based interactions (e.g., triggers and buttons on controllers), and hand-based interactions (e.g., touchable and grabbable) at the same time, so the streamers can intuitively interact with the chat UIs without efforts. Alongside this, avenues of research in information retrieval such as text summarization could generate streaming chat digest for the streamers and send tailored notifications to the streamers for preset events (e.g., new donations or new followers occur) that can also improve the communication and builder stronger connections between streamers and viewers [13, 22, 35, 39].

Visualize and spatialize chat emotions (B2). In our study, we observed numerous instances where viewers formed strong emotional connections with streamers. However, their means of expressing these emotions were limited to sending stickers, emojis, and phrases like "love you" in the chat. Since these expressions are text-based, the VR streamers cannot instantly connect with the viewers' intense emotions at the exact moment. Instead, they have to wait until they check the streaming texts, which significantly hampers the emotional connection between streamers and viewers. To overcome this limitation, prior research has investigated techniques such as summarization, aggregation, and labeling of live streaming content and chats to improve the experience for both streamers and viewers by facilitating content navigation and digestion [21, 35, 39]. Furthermore, the potential of future tools lies in their ability to enhance VR streaming experiences through improved detection and communication of expressed emotions within streaming chats.

By applying sentiment analysis to chat texts, additional insights can be provided to streamers, empowering them to respond adeptly and cultivate their communities with better sensitivity. For example, the chat panel and the VR environment could slightly change their ambiance color to reflect the detected emotions (e.g., tint with light yellow for joy). In addition, the emotions could be visualized as interactable objects popping up around the streamers in VR (e.g., a group of smiling faces around the streamers for happiness). In this way, the streamers and the viewers can more easily build stronger emotional connections and make the streaming experience more engaging.

5.1.3 ***Promote Broader Interactions between Displays and Space.*** Section 4.4 highlights the presence of "undesigned scenes and experiences" in VR streaming. These scenes refer to frustrating static views that offer viewers little to do, often occurring during game loading or when streamers are briefly absent and reveal their streaming space. As highlighted by Glickman et al., an engaging streaming experience requires mutual awareness and co-presence of streamers and viewers, facilitating communication and relationship building between them, and avoiding to overwhelm viewers' screens [23]. Given the prevalent use of interactive devices such as phones, tablets, and smartwatches alongside main screens, an opportunity arises to take full advantages of these additional interactive channels for fostering co-present interactions between streamers and viewers. Future tools should encourage cross-device and spatial interactions to fully utilize the available space and devices of both streamers and viewers, unlocking the potential for a more immersive and engaging streaming experience.

Digitize and customize streaming space (C1). Blurring the line between physical and virtual environments can enhance viewer engagement in streaming [5, 40]. Streamers can digitize and upload their streaming space using 3D scanning and scene reconstruction. This digitalized streaming space can empower streamers to express themselves more effectively and provides viewers with additional spatial context, enhancing their overall streaming experience. The digitized streaming space does not have to be an exact copy of the physical streaming environment. Instead, it should be allowed to be further customized by the streamers (e.g., adding interactable 3D objects and changing materials and textures) to better express the streamers' characteristics and enhance their personal branding. For the viewers, they can freely explore the digitized streaming space while watching the streaming, particularly when the games are loading or the streamers are absent, by walking around the room and playing with the preset interactable objects.

Enable multiple input and output devices (C2). Interactive and rich multimedia elements not only enhance interactions between streamers and viewers but also enrich the overall streaming experience [13, 21, 29, 35, 39]. In addition to providing more interactive multimedia, future tools have the potential to integrate multiple sources as input and output devices like smartwatches, alongside the main display. This integration would enable additional communication channels and create more opportunities for interactions between streamers and viewers, further enhancing the streaming experience. For example, while watching the streamer singing in VR, the viewers can hold their mobile phones and use them as inputs for showing their hand movements like waving and raising hands in VR for the streamers to see. As another example, when the streamer is in a low health state in games and their VR controllers are vibrating, viewers' smartwatches and mobile phones could also vibrate together with the streamers' controllers. In both scenarios, streamers and viewers can have a stronger emotional connection during the streaming, which also unlocks the potential of streaming experience that were previously unexplored and not intentionally designed.

5.2 Methodological Discussion

5.2.1 Reflection on Analysing Streaming Videos. As Anthony pointed out, video data is an important source for observing users' behaviors as they generate content in applications [6]. Recent studies increasingly employ user-generated videos as a data source. However, these studies often use non-interactive videos from platforms like YouTube rather than live streaming sessions. Meanwhile, Harpsted et al. performed an extensive investigation of streaming video research, which highlighted a shift in research focus from streaming platforms to streamers and viewers [30]. Our study resonates with this shift and encounters numerous research challenges outlined in their work, such as effectively processing lengthy streaming videos and analyzing the complex multimodal interactions between streamers and viewers. We want to share our learned lessons for future researchers studying streaming videos.

Analyzing streaming videos poses a significant challenge due to their considerably longer duration compared to typical YouTube videos. Previous research has shown that the average lengths of YouTube videos are typically limited to several minutes (e.g., 127 seconds [6], 3.56 minutes [34], 15.83 seconds [15]). In contrast, streaming videos are usually at least 20 minutes long and can last more than 10 hours. In our study, the shortest video we analyzed is 40 minutes. This creates a huge barrier for coding the full videos if watching them second by second multiple times. To address this challenge, we use an approach we call *transcribe-and-index*. We first recorded the videos and used an auto-transcription service to generate the transcripts of the videos, and then processed these transcripts to derive our codes, leveraging the indexed important points from the transcripts in the videos. This approach allowed us to efficiently navigate and analyze the videos without the need for extensive manual observation.

One question that can be posed is how effective this transcribe-and-index approach is in capturing phenomena. We believe that it is a promising approach for other researchers to adopt. In our study, all streaming videos were fully watched, second by second, by at least one researcher. The method appears to capture most important codes, but has two limitations. It cannot capture soundless interactions (e.g., streamers' body language, on-screen notifications) or fully capture chat conversations. To overcome this, we fully watch each video and trace relevant chat conversations when coding transcripts.

Streaming is gaining increasing popularity globally in recent years. Due to the unique characteristics of streaming videos, as Blythe commented on the limitation of using pure qualitative and quantitative for analysis user generated videos [10], there may not be enough study on the research methods and tools to fully utilize streaming videos as user generated multi-dimension data. As studies continue to focus on processing long-duration streaming videos [21, 39], we expect more future research to enhance the simultaneous integration of text, audio, and video, while streamlining the analysis process through innovative visualization and interactive techniques.

5.2.2 Study Limitations and Future Work. Our study has several limitations. First, our sampling method for choosing the 12 streamers may be biased, as four streamers per game cannot fully represent the range of streaming styles in any specific games. Additionally, our sampling may have favored mature streamers who have more recorded videos and are more likely to be recommended by platforms like Twitch. Thus, our analysis may not fully reflect the behaviors, challenges, and frustrations of less popular streamers. Future work can particularly focus on less popular streamers and abstract their unique challenges for VR streaming. Second, gathering streaming videos posed difficulties due to platforms' limited storage policy. For example, Twitch stores videos for only 14 days for non-prime streamers, making it challenging to find multiple videos for less followed streamers who may not be prime subscribers. Moreover, some streamers delete their past videos, making it nearly impossible to trace streamers' development over time to fully understand how their streaming style evolves in our study. Lastly, relying solely on coded streaming videos may not fully uncover streamers' motivations and rationales for their observed behaviors. Future research could extend our study by triangulating the research data we generate with interviews and surveys to gain deeper insights into streamers' experiences and perspectives.

While streaming has been extensively studied, VR streaming is a relatively new field that offers unique characteristics with the potential to expand our knowledge in both VR and streaming. Recent studies have already observed the impact of non-lab environments and various viewing perspectives on the user experience in VR [15, 18]. VR streaming offers valuable resources for future studies to investigate VR usage and user interactions in real-world settings. Additionally, the growing interest in virtual avatars and virtual spaces for self expressing [8, 9, 12, 40] necessitates research into how VR streaming can serve as an immersive platform for streamers to experience, explore and express their identities alongside their communities. The immersive nature of VR can also amplify discomfort when being observed and recorded by unfamiliar individuals, particularly for those valuing personal space and privacy, underscoring the need for further investigation on privacy and personal data concerns in VR streaming. Finally, VR streaming also requires technological innovation. Similar to streaming studies for viewers' participation [14, 23, 29, 35, 39], exploring how to support similar collaborative VR streaming experience in an asymmetrical condition, with streamers in VR, some viewers in VR, and others using 2D displays, holds promise for future research and development. We believe a closer collaboration between VR game designers and streaming app developers can enhance the streaming experience, enabling innovative design possibilities for game mechanics and viewer interactions tailored for streaming.

6 CONCLUSION

In this paper, we present a multi-phase thematic analysis of 34 streaming videos from 12 VR streamers with different levels of experience on Twitch. We identified four main themes and 18 dimensions to capture the common practices and frustrations of VR streaming. Despite technical challenges, streamers have developed various strategies to engage viewers. Our findings suggest design implications for future VR streaming tools, encouraging collaboration between game designers and streaming app developers. While current tools lack support for diverse interactions, both streamers and viewers desire improved engagement. This underscores the potential benefits of future tools to facilitate interactions between physical and virtual environments, streamers and viewers, and across 2D displays and 3D space.

REFERENCES

[1] [n. d.]. Otter.ai. https://www.otter.ai
[2] [n. d.]. Record & stream yourself from VR with LIV. https://www.liv.tv/
[3] [n. d.]. Twitch.tv - Press Center. https://www.twitch.tv/p/press-center/
[4] [n. d.]. Welcome to Meta | Meta. https://about.facebook.com/meta/
[5] Laurensia Anjani, Terrance Mok, Anthony Tang, Lora Oehlberg, and Wooi Boon Goh. 2020. Why do people watch others eat food? An Empirical Study on the Motivations and Practices of Mukbang Viewers. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1–13. http://doi.org/10.1145/3313831.3376567
[6] Lisa Anthony, YooJin Kim, and Leah Findlater. 2013. Analyzing user-generated youtube videos to understand touchscreen use by people with motor impairments. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1223–1232. http://doi.org/10.1145/2470654.2466158
[7] Verena Biener, Daniel Schneider, Travis Gesslein, Alexander Otte, Bastian Kuth, Per Ola Kristensson, Eyal Ofek, Michel Pahud, and Jens Grubert. 2020. Breaking the Screen: Interaction Across Touchscreen Boundaries in Virtual Reality for Mobile Knowledge Workers. *IEEE Transactions on Visualization and Computer Graphics* 26, 12 (Dec. 2020), 3490–3502. https://doi.org/10.1109/TVCG.2020.3023567 Conference Name: IEEE Transactions on Visualization and Computer Graphics.
[8] Lindsay Blackwell, Nicole Ellison, Natasha Elliott-Deflo, and Raz Schwartz. 2019. Harassment in Social VR: Implications for Design. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. 854–855. https://doi.org/10.1109/VR.

2019.8798165 ISSN: 2642-5254.

[9] Lindsay Blackwell, Jean Hardy, Tawfiq Ammari, Tiffany Veinot, Cliff Lampe, and Sarita Schoenebeck. 2016. LGBT Parents and Social Media: Advocacy, Privacy, and Disclosure during Shifting Social Movements. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. Association for Computing Machinery, New York, NY, USA, 610–622. https://doi.org/10.1145/2858036.2858342

[10] Mark Blythe and Paul Cairns. 2009. Critical methods and user generated content: the iPhone on YouTube. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1467–1476. http://doi.org/10.1145/1518701.1518923

[11] Virginia Braun and Victoria Clarke. 2012. Thematic analysis. In *APA handbook of research methods in psychology, Vol. 2. Research designs: Quantitative, qualitative, neuropsychological, and biological*, H. Cooper, P. M. Camic, D. L. Long, A. T. Panter, D. Rindskopf, and K. J. Sher (Eds.). American Psychological Association, 57–71. https://doi.org/10.1037/13620-004

[12] Matthew Carrasco and Andruid Kerne. 2018. Queer Visibility: Supporting LGBT+ Selective Visibility on Social Media. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3173574.3173824

[13] Di (Laura) Chen, Dustin Freeman, and Ravin Balakrishnan. 2019. Integrating Multimedia Tools to Enrich Interactions in Live Streaming for Language Learning. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3290605.3300668

[14] John Joon Young Chung, Hijung Valentina Shin, Haijun Xia, Li-yi Wei, and Rubaiat Habib Kazi. 2021. Beyond Show of Hands: Engaging Viewers via Expressive and Scalable Visual Communication in Live Streaming. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. Number 109. Association for Computing Machinery, New York, NY, USA, 1–14. http://doi.org/10.1145/3411764.3445419

[15] Emily Dao, Andreea Muresan, Kasper Hornbæk, and Jarrod Knibbe. 2021. Bad Breakdowns, Useful Seams, and Face Slapping: Analysis of VR Fails on YouTube. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. Number 526. Association for Computing Machinery, New York, NY, USA, 1–14. http://doi.org/10.1145/3411764.3445435

[16] Vaibhav Diwanji, Abigail Reed, Arienne Ferchaud, Jonmichael Seibert, Victoria Weinbrecht, and Nicholas Sellers. 2020. Don't just watch, join in: Exploring information behavior and copresence on Twitch. *Computers in Human Behavior* 105 (April 2020), 106221. https://doi.org/10.1016/j.chb.2019.106221

[17] Leah Emerson, Riley Lipinski, Heather Shirey, Theresa Malloy, and Thomas Marrinan. 2021. Enabling Collaborative Interaction with 360° Panoramas between Large-scale Displays and Immersive Headsets. In *2021 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*. 183–188. https://doi.org/10.1109/ISMAR-Adjunct54149.2021.00045

[18] Katharina Emmerich, Andrey Krekhov, Sebastian Cmentowski, and Jens Krueger. 2021. Streaming VR Games to the Broad Audience: A Comparison of the First-Person and Third-Person Perspectives. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. Number 445. Association for Computing Machinery, New York, NY, USA, 1–14. http://doi.org/10.1145/3411764.3445515

[19] Travis Faas, Lynn Dombrowski, Alyson Young, and Andrew D. Miller. 2018. Watch Me Code: Programming Mentorship Communities on Twitch.tv. *Proceedings of the ACM on Human-Computer Interaction* 2, CSCW (Nov. 2018), 50:1–50:18. https://doi.org/10.1145/3274319

[20] Colin Ford, Dan Gardner, Leah Elaine Horgan, Calvin Liu, a. m. tsaasan, Bonnie Nardi, and Jordan Rickman. 2017. Chat Speed OP PogChamp: Practices of Coherence in Massive Twitch Chat. In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA '17)*. Association for Computing Machinery, New York, NY, USA, 858–871. https://doi.org/10.1145/3027063.3052765

[21] C. Ailie Fraser, Joy O. Kim, Hijung Valentina Shin, Joel Brandt, and Mira Dontcheva. 2020. Temporal Segmentation of Creative Live Streams. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3313831.3376437

[22] C. Ailie Fraser, Joy O. Kim, Alison Thornsberry, Scott Klemmer, and Mira Dontcheva. 2019. Sharing the Studio: How Creative Livestreaming can Inspire, Educate, and Engage. In *Proceedings of the 2019 on Creativity and Cognition (C&C '19)*. Association for Computing Machinery, New York, NY, USA, 144–155. https://doi.org/10.1145/3325480.3325485

[23] Seth Glickman, Nathan McKenzie, Joseph Seering, Rachel Moeller, and Jessica Hammer. 2018. Design Challenges for Livestreamed Audience Participation Games. In *Proceedings of the 2018 Annual Symposium on Computer-Human Interaction in Play (CHI PLAY '18)*. Association for Computing Machinery, New York, NY, USA, 187–199. https://doi.org/10.1145/3242671.3242708

[24] Daniel Gros, Brigitta Wanner, Anna Hackenholt, Piotr Zawadzki, and Kathrin Knautz. 2017. World of Streaming. Motivation and Gratification on Twitch. In *Social Computing and Social Media. Human Behavior (Lecture Notes in*

Computer Science), Gabriele Meiselwitz (Ed.). Springer International Publishing, Cham, 44–57. https://doi.org/10.1007/978-3-319-58559-8_5

[25] Jan Gugenheimer, Evgeny Stemasov, Julian Frommel, and Enrico Rukzio. 2017. ShareVR: Enabling Co-Located Experiences for Virtual Reality between HMD and Non-HMD Users. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. Association for Computing Machinery, New York, NY, USA, 4021–4033. https://doi.org/10.1145/3025453.3025683

[26] Lassi Haaranen. 2017. Programming as a Performance: Live-streaming and Its Implications for Computer Science Education. In *Proceedings of the 2017 ACM Conference on Innovation and Technology in Computer Science Education (ITiCSE '17)*. Association for Computing Machinery, New York, NY, USA, 353–358. https://doi.org/10.1145/3059009.3059035

[27] Oliver L. Haimson and John C. Tang. 2017. What Makes Live Events Engaging on Facebook Live, Periscope, and Snapchat. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 48–60. http://doi.org/10.1145/3025453.3025642

[28] William A. Hamilton, Oliver Garretson, and Andruid Kerne. 2014. Streaming on twitch: fostering participatory communities of play within live mixed media. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '14)*. Association for Computing Machinery, New York, NY, USA, 1315–1324. https://doi.org/10.1145/2556288.2557048

[29] William A. Hamilton, Nic Lupfer, Nicolas Botello, Tyler Tesch, Alex Stacy, Jeremy Merrill, Blake Williford, Frank R. Bentley, and Andruid Kerne. 2018. Collaborative Live Media Curation: Shared Context for Participation in Online Learning. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3173574.3174129

[30] Erik Harpstead, Juan Sebastian Rios, Joseph Seering, and Jessica Hammer. 2019. Toward a Twitch Research Toolkit: A Systematic Review of Approaches to Research on Game Streaming. In *Proceedings of the Annual Symposium on Computer-Human Interaction in Play (CHI PLAY '19)*. Association for Computing Machinery, New York, NY, USA, 111–119. https://doi.org/10.1145/3311350.3347149

[31] Jeremy Hartmann and Daniel Vogel. 2022. Enhanced Videogame Livestreaming by Reconstructing an Interactive 3D Game View for Spectators. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems (CHI '22)*. Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3491102.3517521

[32] Juan Pablo Hourcade, Sarah L. Mascher, David Wu, and Luiza Pantoja. 2015. Look, My Baby Is Using an iPad! An Analysis of YouTube Videos of Infants and Toddlers Using Tablets. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1915–1924. http://doi.org/10.1145/2702123.2702266

[33] Marcel Jonas, Steven Said, Daniel Yu, Chris Aiello, Nicholas Furlo, and Douglas Zytko. 2019. Towards a Taxonomy of Social VR Application Design. In *Extended Abstracts of the Annual Symposium on Computer-Human Interaction in Play Companion Extended Abstracts (CHI PLAY '19 Extended Abstracts)*. Association for Computing Machinery, New York, NY, USA, 437–444. https://doi.org/10.1145/3341215.3356271

[34] Aida Komkaite, Liga Lavrinovica, Maria Vraka, and Mikael B. Skov. 2019. Underneath the Skin: An Analysis of YouTube Videos to Understand Insertable Device Interaction. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1–12. http://doi.org/10.1145/3290605.3300444

[35] Pascal Lessel, Alexander Vielhauer, and Antonio Krüger. 2017. Expanding Video Game Live-Streams with Enhanced Communication Channels: A Case Study. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1571–1576. http://doi.org/10.1145/3025453.3025708

[36] Zhicong Lu. 2019. Live streaming in China for sharing knowledge and promoting intangible cultural heritage. *Interactions* 27, 1 (Dec. 2019), 58–63. https://doi.org/10.1145/3373145

[37] Zhicong Lu, Michelle Annett, Mingming Fan, and Daniel Wigdor. 2019. "I feel it is my responsibility to stream": Streaming and Engaging with Intangible Cultural Heritage through Livestreaming. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1–14. http://doi.org/10.1145/3290605.3300459

[38] Zhicong Lu, Michelle Annett, and Daniel Wigdor. 2019. Vicariously Experiencing it all Without Going Outside: A Study of Outdoor Livestreaming in China. *Proceedings of the ACM on Human-Computer Interaction* 3, CSCW (Nov. 2019), 25:1–25:28. https://doi.org/10.1145/3359127

[39] Zhicong Lu, Seongkook Heo, and Daniel J. Wigdor. 2018. StreamWiki: Enabling Viewers of Knowledge Sharing Live Streams to Collaboratively Generate Archival Documentation for Effective In-Stream and Post Hoc Learning. *Proceedings of the ACM on Human-Computer Interaction* 2, CSCW (Nov. 2018), 112:1–112:26. https://doi.org/10.1145/3274381

[40] Zhicong Lu, Chenxinran Shen, Jiannan Li, Hong Shen, and Daniel Wigdor. 2021. More Kawaii than a Real-Person Live Streamer: Understanding How the Otaku Community Engages with and Perceives Virtual YouTubers. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. Number 137. Association for Computing Machinery, New York, NY, USA, 1–14. http://doi.org/10.1145/3411764.3445660

[41] Zhicong Lu, Haijun Xia, Seongkook Heo, and Daniel Wigdor. 2018. You Watch, You Give, and You Engage: A Study of Live Streaming Practices in China. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1–13. http://doi.org/10.1145/3173574.3174040

[42] Deirdre Lyons, Stephen G. Butchko, Jason Moore, Brendan Bradley, and Tanya Leal Soto. 2021. Live Performance in VR: Live performance in virtual reality by creators from different metaverses discuss the challenges and advantages of performance in this new storytelling platform.. In *Special Interest Group on Computer Graphics and Interactive Techniques Conference (SIGGRAPH '21 Panels)*. Association for Computing Machinery, New York, NY, USA, 1–2. https://doi.org/10.1145/3450617.3464495

[43] Raquel Macedo, Nuno Correia, and Teresa Romão. 2019. Paralympic VR Game: Immersive Game using Virtual Reality and Video. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*. ACM, Glasgow Scotland Uk, 1–6. https://doi.org/10.1145/3290607.3312938

[44] Silja Martikainen, Valtteri Wikström, Mari Falcon, and Katri Saarikivi. 2019. Collaboration Face-to-Face and in Virtual Reality - Empathy, Social Closeness, and Task Load. In *Conference Companion Publication of the 2019 on Computer Supported Cooperative Work and Social Computing*. ACM, Austin TX USA, 299–303. https://doi.org/10.1145/3311957.3359468

[45] Michael Nebeling, Katy Lewis, Yu-Cheng Chang, Lihan Zhu, Michelle Chung, Piaoyang Wang, and Janet Nebeling. 2020. XRDirector: A Role-Based Collaborative Immersive Authoring System. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1–12. http://doi.org/10.1145/3313831.3376637

[46] Michael Nebeling, Shwetha Rajaram, Liwei Wu, Yifei Cheng, and Jaylin Herskovitz. 2021. XRStudio: A Virtual Production and Live Streaming System for Immersive Instructional Experiences. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. Number 107. Association for Computing Machinery, New York, NY, USA, 1–12. http://doi.org/10.1145/3411764.3445323

[47] Jeni Paay, Jesper Kjeldskov, and Mikael B. Skov. 2015. Connecting in the Kitchen: An Empirical Study of Physical Interactions while Cooking Together at Home. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing (CSCW '15)*. Association for Computing Machinery, New York, NY, USA, 276–287. https://doi.org/10.1145/2675133.2675194

[48] Rui Pan, Lyn Bartram, and Carman Neustaedter. 2016. TwitchViz: A Visualization Tool for Twitch Chatrooms. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems*. ACM, San Jose California USA, 1959–1965. https://doi.org/10.1145/2851581.2892427

[49] Priyanka Pazhayedath, Pedro Belchior, Rafael Prates, Filipe Silveira, Daniel Simões Lopes, Robbe Cools, Augusto Esteves, and Adalberto L. Simeone. 2021. Exploring Bi-Directional Pinpointing Techniques for Cross-Reality Collaboration. In *2021 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*. 264–270. https://doi.org/10.1109/VRW52623.2021.00055

[50] Anthony J. Pellicone and June Ahn. 2017. The Game of Performing Play: Understanding Streaming as Cultural Production. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 4863–4874. http://doi.org/10.1145/3025453.3025854

[51] David Saffo, Caglar Yildirim, Sara Di Bartolomeo, and Cody Dunne. 2020. Crowdsourcing Virtual Reality Experiments using VRChat. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*. ACM, Honolulu HI USA, 1–8. https://doi.org/10.1145/3334480.3382829

[52] John Scott Siri Jr., Hamna Khalid, Luong Nguyen, and Donghee Yvette Wohn. 2018. Screen-viewing Practices in Social Virtual Reality. In *Companion of the 2018 ACM Conference on Computer Supported Cooperative Work and Social Computing*. ACM, Jersey City NJ USA, 173–176. https://doi.org/10.1145/3272973.3274048

[53] Max Sjöblom and Juho Hamari. 2017. Why do people watch others play video games? An empirical study on the motivations of Twitch users. *Computers in Human Behavior* 75 (Oct. 2017), 985–996. https://doi.org/10.1016/j.chb.2016.10.019

[54] Joshua Suttor, Julian Marin, Evan Verbus, and Meng Su. 2019. Implement AI Service into VR Training. In *Proceedings of the 2019 2nd International Conference on Signal Processing and Machine Learning*. ACM, Hangzhou China, 114–121. https://doi.org/10.1145/3372806.3374909

[55] John C. Tang, Gina Venolia, and Kori M. Inkpen. 2016. Meerkat and Periscope: I Stream, You Stream, Apps Stream for Live Streams. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. Association for Computing Machinery, San Jose, California, USA, 4770–4780. https://doi.org/10.1145/2858036.2858374

[56] T. L. Taylor. 2018. *Watch Me Play: Twitch and the Rise of Game Live Streaming*. Princeton University Press. http://www.jstor.org/stable/j.ctvc77jqw

[57] Balasaravanan Thoravi Kumaravel, Fraser Anderson, George Fitzmaurice, Bjoern Hartmann, and Tovi Grossman. 2019. Loki: Facilitating Remote Instruction of Physical Tasks Using Bi-Directional Mixed-Reality Telepresence. In *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology (UIST '19)*. Association for Computing Machinery, New York, NY, USA, 161–174. https://doi.org/10.1145/3332165.3347872

[58] Balasaravanan Thoravi Kumaravel, Cuong Nguyen, Stephen DiVerdi, and Björn Hartmann. 2019. TutoriVR: A Video-Based Tutorial System for Design Applications in Virtual Reality. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1–12. http://doi.org/10.1145/3290605.3300514

[59] Balasaravanan Thoravi Kumaravel, Cuong Nguyen, Stephen DiVerdi, and Bjoern Hartmann. 2020. TransceiVR: Bridging Asymmetrical Communication Between VR Users and External Collaborators. In *Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology (UIST '20)*. Association for Computing Machinery, New York, NY, USA, 182–195. https://doi.org/10.1145/3379337.3415827

[60] Balasaravanan Thoravi Kumaravel and Andrew D Wilson. 2022. DreamStream: Immersive and Interactive Spectating in VR. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems (CHI '22)*. Association for Computing Machinery, New York, NY, USA, 1–17. https://doi.org/10.1145/3491102.3517508

[61] Oculus VR. [n. d.]. Mixed Reality with Passthrough. https://developer.oculus.com/blog/mixed-reality-with-passthrough/

[62] Frederik Winther, Linoj Ravindran, Kasper Paabol Svendsen, and Tiare Feuchtner. 2020. Design and Evaluation of a VR Training Simulation for Pump Maintenance Based on a Use Case at Grundfos. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE, Atlanta, GA, USA, 738–746. https://doi.org/10.1109/VR46266.2020.00097

[63] Donghee Yvette Wohn. 2019. Volunteer Moderators in Twitch Micro Communities: How They Get Involved, the Roles They Play, and the Emotional Labor They Experience. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3290605.3300390

[64] Donghee Yvette Wohn, Guo Freeman, and Caitlin McLaughlin. 2018. Explaining Viewers' Emotional, Instrumental, and Financial Support Provision for Live Streamers. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3173574.3174048

[65] Donghee Yvette Wohn, Peter Jough, Peter Eskander, John Scott Siri, Masaho Shimobayashi, and Pradnya Desai. 2019. Understanding Digital Patronage: Why Do People Subscribe to Streamers on Twitch?. In *Proceedings of the Annual Symposium on Computer-Human Interaction in Play (CHI PLAY '19)*. Association for Computing Machinery, New York, NY, USA, 99–110. https://doi.org/10.1145/3311350.3347160

[66] Saelyne Yang, Changyoon Lee, Hijung Valentina Shin, and Juho Kim. 2020. Snapstream: Snapshot-based Interaction in Live Streaming for Visual Art. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1–12. http://doi.org/10.1145/3313831.3376390

[67] Enes Yigitbas, Ivan Jovanovikj, Janis Scholand, and Gregor Engels. 2020. VR Training for Warehouse Management. In *26th ACM Symposium on Virtual Reality Software and Technology*. ACM, Virtual Event Canada, 1–3. https://doi.org/10.1145/3385956.3422106

[68] Soojeong Yoo, Phillip Gough, and Judy Kay. 2020. Embedding a VR Game Studio in a Sedentary Workplace: Use, Experience and Exercise Benefits. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. ACM, Honolulu HI USA, 1–14. https://doi.org/10.1145/3313831.3376371

[69] Samaneh Zamanifard and Guo Freeman. 2019. "The Togetherness that We Crave": Experiencing Social VR in Long Distance Relationships. In *Conference Companion Publication of the 2019 on Computer Supported Cooperative Work and Social Computing (CSCW '19)*. Association for Computing Machinery, New York, NY, USA, 438–442. https://doi.org/10.1145/3311957.3359453

[70] Nima Zargham, Michael Bonfert, Georg Volkmar, Robert Porzel, and Rainer Malaka. 2020. Smells Like Team Spirit: Investigating the Player Experience with Multiple Interlocutors in a VR Game. In *Extended Abstracts of the 2020 Annual Symposium on Computer-Human Interaction in Play*. ACM, Virtual Event Canada, 408–412. https://doi.org/10.1145/3383668.3419884

[71] Keran Zhao, Yuheng Hu, Yili Hong, and Christopher Westland. 2021. Understanding Characteristics of Popular Streamers on Live Streaming Platforms: Evidence from Twitch.tv. *Journal of the Association for Information Systems* 22, 4 (July 2021), 4. https://doi.org/10.17705/1jais.00689 Num Pages: 4 Place: Atlanta, United States Publisher: Association for Information Systems Section: Articles.

[72] Jilei Zhou, Jing Zhou, Ying Ding, and Hansheng Wang. 2018. The Magic of Danmaku: A Social Interaction Perspective of Gift Sending on Live Streaming Platforms. *SSRN Electronic Journal* (2018). https://doi.org/10.2139/ssrn.3289119

[73] Yu Zhu, Kang Zhu, Qiang Fu, Xilin Chen, Huixing Gong, and Jingyi Yu. 2016. SAVE: shared augmented virtual environment for real-time mixed reality applications. In *Proceedings of the 15th ACM SIGGRAPH Conference on Virtual-Reality Continuum and Its Applications in Industry - Volume 1 (VRCAI '16)*. Association for Computing Machinery, New York, NY, USA, 13–21. https://doi.org/10.1145/3013971.3013979

Received 2023-07-01; accepted 2023-09-22