Designing for Shared Remote Video Experiences

Anders Sundnes Løvlie
asun@itu.dk
IT University of Copenhagen
Copenhagen, Denmark

Lucas Schønrock
lucas.schonrock@gmail.com
IT University of Copenhagen
Copenhagen, Denmark

Jannik Bryld
jannikbryld@gmail.com
IT University of Copenhagen
Copenhagen, Denmark

Louise Barkhuus
barkhuus@itu.dk
IT University of Copenhagen
Copenhagen, Denmark

ABSTRACT

Watching television in the living room is a joyful, traditional activity, often shared with friends and family. While sociality is becoming increasingly distributed as a result of emerging technologies, modern life and geographical distance (and the COVID-19 pandemic) that prevents physical gatherings, the desire to socialize around video content remains. This paper describes a study of sociable viewing experiences enabled by a web-based video sharing application that includes a virtual "remote control" that helps users coordinate content selection and playback, and a drawing feature intended to help facilitate new forms of social interaction with online video. Our results focus on two themes: What modes of communication are needed in order to create a sense of co-presence, and how can we design such system to enable new forms of sociability. Finally, we report user patterns from a public available version of out system. Our study contributes to a further understanding of remote social video and television watching and illustrate how we can better design for the social watching.

CCS CONCEPTS

• **Human-centered computing** → *Interactive systems and tools*;
• **Information systems** → *Web conferencing*.

KEYWORDS

Video, presence, social watching, user experience, covid-19, social distance

ACM Reference Format:
Anders Sundnes Løvlie, Lucas Schønrock, Jannik Bryld, and Louise Barkhuus. 2018. Designing for Shared Remote Video Experiences. In *Woodstock '18: ACM Symposium on Neural Gaze Detection, June 03–05, 2018, Woodstock, NY*. ACM, New York, NY, USA, 11 pages. https://doi.org/10.1145/1122445.1122456

Woodstock '18, June 03–05, 2018, Woodstock, NY

ACM ISBN 978-1-4503-XXXX-X/18/06…$15.00
https://doi.org/10.1145/1122445.1122456

1 INTRODUCTION

Spurred on by the COVID-19 pandemic, as well as broader societal and technological trends, people have increasingly explored ways to carry out social activities through online video connections, from virtual cocktail hour to virtual cooking courses. While communication technologies might not have reached the level of "Beyond Being there" [16] as Hollan and Stornetta speculated upon already in 1992, our social lives have at least been augmented with more sophisticated remote opportunities. This paper explores the idea of social TV-watching online: To re-stage the experience of getting together to watch a film or a TV show, using an online connection instead of getting physically together in someone's living room.

Television viewing has often been perceived as an isolated, anti-social activity. However, scholars have also pointed to video media as opportunities for sociability [21, 29]. The success of video-on-demand services such as Netflix, Disney+, YouTube and television channels' own services (e.g. BBC's iPlayer) has led to significant changes in viewing habits, with TV content increasingly being watched time shifted, mostly through the Internet [8]. Still, limited research has looked at sociability around remote co-watching, further than experimental co-watching systems and studies of shared video consumption. While experimental systems for interacting remotely around content watching have been developed and trialled, many of these systems have focused on traditional, non-time shifted television watching and application specific systems [7]. Recent improvements in network speed and web technologies have made it increasingly feasible to launch web-based systems for remote co-watching experiences and several systems are already commercially available.

In this paper we present a web-based system that we developed through several design iterations and three rounds of user testing. Our goal was to develop a web-based prototype that facilitates distributed shared video experiences, in order to explore potential interaction and possible features that can contribute to a sociable experience. While similar services have also been developed by others, our study is the first academic study building and exploring a fully developed video sharing service that uses web-technologies and works straight from a browser. Initially this web-based service was built as a research project and enabled us to explore sociability and viewing experiences among a test audience. At a later stage the service was developed into a commercial product and deployed as a feature in the online video player of a national broadcaster, enabling us to present"in-the-wild" data from a general audience.

Our broader research goal was to explore the new forms of social experiences made possible by a system that offers users rich channels of communication with each other, while watching synchronous streaming video content - analogous to the living room situation.

Our initial focus was on two themes in particular: What modes of communication were needed in order to create a sense of co-presence? And how could the system help the users coordinate content selection and playback? Through the process, a third theme emerged: What new forms of sociability might be enabled by the system, and how could the system be designed to facilitate these interactions? After developing a prototype named "LetsWatch", we conducted a qualitative evaluation that led to a better understanding of the potential uses and challenges in video sharing systems. Finally we partnered with a national public broadcaster to develop and release a version which currently is available for online watching on their Video-on-Demand web service. Our findings include usage data from the first four months of full public availability, allowing us to extend the insights from initial prototyping and user testing with data from use by a general audience in their everyday context. This paper contributes by describing a novel synchronous video sharing system, including its design process, as well as observations from early experiences and deployment to a broader public service.

2 RELATED WORK

We begin by looking at remote interaction between colleagues, friends and family in a broad sense. Much of this research is rooted in the field of Computer Supported Cooperative Work (CSCW), and while it dates back to the early nineties, it is still relevant to understand the foundation of remote social interaction. Following that we look at what television viewers do while they are watching television in groups, both when they are physically co-present and geographically distributed. Past research has focused mostly on investigating how viewers' social demographics affect their choice of programs [12], rather than how they behave when watching video content, and less is known, particular in real settings, about interaction patterns in front of the television screen [10].

2.1 Social Remote Interaction

Sharing a screen for collaboration has been a core research subject for decades, initially within the area of Groupware, but mostly now as part of the CSCW area. One clear differentiation is between co-location and remote collaboration: where some systems are designed for co-location [18], others are explicitly designed and developed for remote collaboration, for example where edits on one screen shows up real-time on a remote colleague's screen [14, 27].

One major distinction between these systems and the one we present in this paper, is that the majority of collaborative remote systems mentioned above were designed for work purposes, while our system is designed for leisure and social activities. When considering sociability among family and friend, research has focused on for example remote storybook reading by family members [28], remote video chat [1], and presence displays [9]. Unique systems such as the "messaging kettle" [6] and picture frame interaction [19] have also been designed, prototyped and tested, mostly focusing on remote everyday interaction with family members.

Several of these previous systems have similar goals to our leisure-based prototype - such as providing remote interaction with friends and family - but relating to other areas of leisure. We therefore now turn to research explicitly focused on socializing in relation to video based entertainment.

2.2 Understanding Video-Mediated Sociability

People often use movie nights (an evening where friends and family gather together to watch television or video content) as an occasion for sociality [12]. Liu et al. describe these collocated viewing parties and the interactions that occur as: "[...] people gather, such as on a couch, for a social experience. They can co-watch, exchange comments, and interact with each other. Such synchronized group interactions tend to foster relations and social capital among the participants." [20, p. 3] Interactions between viewers are tightly interwoven with the context of the video they are watching and it is common to talk about the content while the video is playing. In an empirical study, Ducheneau et al. [10] analyzed the unspoken interaction rules that allow participants to simultaneously socialize with each other around the TV and follow the ongoing program with enough attention to understand what is happening. The study found that this socializing behaviour could continue also when participants were placed in separate rooms with an audio connection, suggesting that: "groups can socialize remotely while watching TV using a simple, always on audio channel" [10, p. 6].

Williams et al. found that social communication was highlighted as potentially beneficial in distributed shared video experiences and recommended that "high-quality videoconferencing capabilities should be brought into the domestic environment and made to operate on high definition TVs over contended broadband networks" [33, p. 27]. Geerts et al. [13] conducted a similar study on distributed shared video experiences and found that synchronization of video content is critical for the quality of the shared video experience. Additionally, they found that viewers should be able to interact with the video content, be able to pause and play to let participants be able to talk about specific frames in a video.

2.3 The Internet as Video Media Consumption Facilitator

The practice of using Internet as a tool for video content streaming emerged in the early 2000s as a result of faster broadband Internet as well as cheaper desktop computers [3, 4]. Up until 2017, more people had cable subscriptions but by then the major streaming service Netflix, had surpassed in terms of numbers of subscriptions [25]. By 2021, 26% of all content in the US is watched through streaming services and this trend is expected to continue in the future.

Historically, television viewing was a synchronous experience. Every TV receiver would receive the same broadcast signal and play it roughly synchronous to everybody else, watching the same channel. This means that personal interaction around content would either be direct, synchronous interaction or semi-asynchronous 'water-cooler talk' about immediate past content ("Did you see the last episode of X last night?"). Accessing content through the Internet has introduced new forms of asynchronous viewing as well as new forms of indirect sociability. Now, most content on the

web is asynchronous, so-called time shifted or on-demand content, enabling viewers to choose precisely what, when and where to watch content.

While asynchronous video and indirect communication may increase commonality, there is a decreasing option for co-presence and synchronicity, which is essential for actual shared video experiences [22]. In a study exploring shared video experiences, Liu et al. [20] implemented a system that automatically synchronized video playback between a group of geographically separated viewers and utilized a synchronous IM (instant messaging) system to help geographically separated viewers maintain direct sociability. This study found that while it is possible to encourage video-mediated direct sociability on the Internet, despite its asynchronous nature, it is often not sufficient to rely on textual communication alone to establish co-presence. Finally, the study found that the current system's limit of just two concurrent viewers should be increased to allow for more viewers.

Relatedly, Macaranas et al. [22] conducted an extensive study in which geographically separated viewers were asked to watch television while communicating through the web-based service Skype to simulate a distributed shared video experience. The study found that remotely watching a video together is a desirable, fun activity that augments social communication and reinforces social connection and presence to a degree comparable with the shared TV experience as previously defined. The study found that while audio between viewers is fundamental for communication, video actually appears to play a role in creating a social presence. Some viewers expressed that the shared video experience did not feel natural and was not as pleasant as watching in person. Macaranas et al. noted that technical complexity and content synchronization are the two primary barriers for remote sociable video watching, arguing that technical problems can drastically decrease the effectiveness of the overall experience [22]. Their overall conclusion was that shared video experiences do in fact encourage video-mediated sociability, as long as it allows close ties to strengthen and as long as the experiences are complementary to "simply conversing together" [22, p. 18].

From a more theoretical perspective, Cesar and Geerts [7] define a framework for shared television viewing as consisting of the following four categories:

- Content selection and sharing: Information by other peers is used for making appropriate decisions on what to watch.
- Communication: Direct communication via chat, audio, or video with other peers while watching television content.
- Community building: Commenting about a television program with a large community of viewers.
- Status update: Making available to others what you are currently watching.

While Cesar and Geerts argue that the framework is sufficient to describe present and past shared television viewing solutions, it is aimed at social-media inspired solutions where viewers might be separated in time and in space, and thus not facilitating the same kind of direct sociability that occurs during traditional viewing parties. Still, the framework presents an interesting perspective that was used as inspiration when developing the prototype.

2.4 Social Watching Built for Research

Previous social watching technologies and services have been developed and tested both from research and development perspectives, as well as for commercial purposes. An early prototype, "AmigoTV", was used to investigate difference in preference of voice interaction or "icon interaction" (reactions) in a remote watching situation. The study concluded that remote voice interaction had great potential, similarly to face-to-face watching where discussions are common [2]. Harboe et al. [15] developed a set of ambient displays that allowed users to see which television shows their friends and family were currently watching as well as support exchange of short messages within the TV-viewing system. Their aim was to provide ambient social presence for remote people. Collabora-TV is another prototype system where users can annotate their video content for friends to watch at another time [26].

Recent studies have explored entirely different approaches to co-watching and socialising remotely. McGill et al. [24] use smart TVs and VR head-mounted displays to explore remote co-watching for pairs, with a variety of content types. The same authors have explored control of content in multi-user systems, concluding in favor of systems that give control to one user at a time rather than all users simultaneously [23]. Conversely, Sun et al. [31] studied co-watching of short-form video content, and offer design recommendations which emphasize facilitating group interactions both in searching for content and agreeing on what to watch. Feltwell et al. [11] explore the potential for second-screening to facilitate critical reflection, through a mobile application that enables co-voting and live textual tagging. In terms of direct screen interaction, many systems have been experimented with, such as interactive video streams at live performances [5] and interactive live choreographer applications [30]. Other related systems include prototypes facilitating interactive overlay elements in a video, and interactive links between video and a corresponding digital game [32].

In terms of existing commercial systems, until recently few systems had been developed to support social interactions for television as well as web-based video services. None of them have been used for in-depth studies in social watching, which is why we devote a separate section to these.

3 STATE OF THE ART

3.1 Commercial Products

In the following, we look at existing web services that support *shared remote* video experiences - meaning that they support multiple simultaneous users and that both the playback of video content and communication medium must be synchronized between users, facilitating synchronous sociability. Mobile applications not usable in a browser, such as *AirTime*[1], have been considered but are likewise excluded from this review as the focus of this project is on web applications.

Table 1 presents seven commercial products that lie close to the problem domain. These products exist either as stand-alone applications or as overlay applications, which are extensions or plugins to already existing applications and provide extra functionality such as

[1] https://www.airtime.com/

Table 1: Identified state-of-the-art commercial products that facilitate shared remote video experiences. "Local stream" refers to the possibility of sharing video files or screen sharing.

Service	Communication	Content selection	Content playback
Facebook Watch Party	Text	Facebook content	Only host control
Kast Party	Text, speech, webcam	Local stream	Users control own broadcast
Watch2gether	Text, speech, webcam	Free-to-watch services	Equal control
TwoSeven	Text, speech, webcam	Netflix, other subscription services, free-to-watch services and local stream	Equal control (or only host)
MetaStream	Text	Netflix, other subscription services and free-to-watch services	Equal control
Netflix Party	Text	Netflix	Equal control (or only host)
Scener	Text, speech, webcam	Netflix	Equal control

video playback synchronization. The stand-alone applications identified are: *Facebook watch party*, *Kast*, *Watch2Gether*, *TwoSeven* and *Metastream*. The overlay applications identified are *Netflix Party* and *Scener*. Table 1 shows a summary of the services categorized in terms an adapted version of [7]'s framework. These categories include the social interactions, video content, and control schemas that they facilitate. Though many of the services are quite similar, differences are found in how users can socialize and how video selection and playback are achieved, as well as how the different experiences look and feel.

3.1.1 Stand-Alone Applications. In 2018, Facebook released a new feature called Facebook watch party[2], which allows users of Facebook to initiate a *watch party* when creating a new post in a group. Users that view the post can then choose to participate in the party. All participating users of a watch party are depicted by their profile picture in the lower-left corner. The video shown is synchronized between all users. The users can also chat together through a text chat. There are no options, however, for users to communicate via speech or webcam. The user that initiated the watch party is denoted *host* and has full control over video selection, pausing/resuming, and seeking of the active video. Other users may still suggest videos to watch, which the host subsequently may choose to approve or disregard. The selectable videos for a party are limited to public videos hosted on Facebook, and it is not possible to select external video sources of any kind.

Another relevant web application is *Kast*[3], which also facilitates watch parties that much resembles the aforementioned Facebook watch parties, although with a few significant differences in terms of functionality. In a Kast watch party users can also communicate together through speech. Another fundamental difference between the two services is that in Kast, the host (party creator) can grant permissions to other party members, so they also can share a video stream. This means that it is possible for everyone to share a video, and thus there may be multiple videos shown in a party at any one time. The video content limitation in Facebook watch parties is not present in a Kast watch party, as users never share any specific online video, but instead chooses to screen-share or share their webcam feed. This way, users can share nearly any kind of video content, but it simultaneously hinders any kind of involvement from participants in terms of being able to play, pause, and seek collaboratively in a video. Another limitation of this design is that a user cannot both broadcast their webcam and a video simultaneously (because the feed shows up in the same panel), but instead have to choose between either of the two.

Watch2Gether[4] is a service that facilitates an experience close to that of Facebook Watch Party and Kast but differentiates in what video content is available for playback. Watch2Gether, in fact, directly interfaces with a large number of public video services, including Youtube, *Vimeo*, and *Twitch*. This type of public free-to-watch video content was found to be rather common in the investigated products, also reflected in Table 1. Watch2Gether has *rooms*, which work much like watch parties, where other users can be invited through a unique URL. In this room, users can communicate via webcam, speech, and a text chat, while watching video together, unlike in Facebook watch party and Kast. Everyone can equally pause, play, search for, and queue content amongst all the supported services. If a video is put into the playlist queue the next video in the queue will automatically play after a video has ended.

The two last stand-alone web applications investigated are *TwoSeven*[5] and *MetaStream*[6]. Functionality-wise they are quite similar to the three prior services. However, they provide a different take on what type of video content can be viewed. While the other services only facilitate sharing of free-to-watch services, TwoSeven and MetaStream both utilize *Chrome Browser* technology (Chrome extensions) to embed video content providers into the platform. As such, users can log in to account-restricted and copyright-protected content providers such as Netflix, directly within both Metastream or Twoseven, and can then share the video player of, for example, Netflix. It is, however, required to install a third-party chrome extension (created by the services themselves) for this to function.

3.1.2 Overlay Applications. Netflix Party[7] and *Scener*[8] are two products that, like Metastream and Twoseven, require the installation of a third-party chrome extension to function. In fact, Netflix

[2]https://www.facebook.com/help/1681245065258554 (Visited 08-04-2020)
[3]https://kastapp.co/ (Visited 02-04-2020)
[4]https://watch2gether.com/ (Visited 08-04-2020)
[5]https://twoseven.xyz/ (Visited 08-04-2020)
[6]https://getmetastream.com/ (Visited 08-04-2020)
[7]https://netflixparty.com/ (Visited 11-04-2020)
[8]https://scener.com/ (Visited 11-04-2020)

Party and Scener do not have a stand-alone web application at all. Instead, they inject (add) synchronization and communication functionality directly onto Netflix in the browser, which is the exclusive content provider for the two applications. As such, both extensions take control of the Netflix video player and synchronizes any action that users take. Users, therefore, have to install the third-party extension, login to Netflix, start playback of a video, and then share a unique link with each other before they can begin watching together. Netflix Party is designed inherently simpler than Scener, as it only provides a possibility for users of a shared video experience to communicate through a text chat. Scener, on the other hand, while providing the same core functionality as Netflix Party, provides many more ways for users to interact through a more sophisticated user interface. Upon having installed the third party chrome extension, users are required to create a profile with email, name, and password and to select a profile image. After this, a new sidebar will appear when watching video content on Netflix, analogous to Netflix Party. This sidebar, however, includes novel features not seen on any of the other services investigated. In particular, Scener provides a friend-list system, allowing users to keep track of which of their friends are on Netflix and what they are watching, also providing an easy way for them to interact and initiate a shared viewing session. All users, including oneself, is represented by a circle in the top right corner, which shows the chosen profile picture or a webcam feed of the participants. Scener, unlike Netflix Party, allows for both a webcam and speech communication in addition to the text chat. In the time since this state of the art review was conducted, some further commercial services have been launched - such as Disney+ Groupwatch and Amazon Prime Video Watch Party, offering features similar to the ones discussed above.

4 DESIGN

Our development strategy was one of user centered design with rapid iterations. The prototype was developed using agile development processes and design sprints, in order to quickly have a product we could test on users [17]. The goal of the multiple design iterations with users was to establish a better understanding of web-based sociable video watching through development of relevant and appropriate functionality. The criteria for development were to support sociability as well as possible in a remote situation, as well as provide a wide set of options for interacting with different video content. To support this iterative process we carried out several small user tests throughout development, in order to enable a rapid redesign process and a flexible design. The test and evaluation methodology is described in the next section.

The design process initially focused in particular on two questions: What modes of communication would be needed in order to create a sense of co-presence? And how could the system help the users coordinate content selection and playback?

Regarding the first of these two questions, iterative prototyping and testing painted a consistent picture that users desired having access to all the communication modalities we could make available - text chat, audio and video (webcam) - in order to communicate in a way that felt natural. Having both audio and webcam channels open allowed users to communicate in a more dynamic and flexible manner, making quick comments during pauses in the dialogue in the video, and even when not directly communicating the webcam feed allowed them to sense the presence of the other users.

Contemplating the second theme, coordination, led us to design a concept called "the virtual living room", which would serve as a private space that users could invite other users into. Through iterations this concept was implemented as two separate screens, one for *video selection* (Figure 1) in which users can search or browse for content, and a second screen for *video playback* (Figure 2). The main features of the video selection screen are the search bar (feature 1 in Figure 1), search results (feature 2), options for watching live TV (feature 3), as well as an option for accessing videos from social media by inserting a link to the video (feature 4).

When a video is chosen, the page changes for all participating users to the video playback screen, which was designed to let the main video fill most of the screen, whereas other interface elements such as webcam feeds and text messages from other users are displayed on the edges. This page shows information about the selected content in feature 6, and a playbar in feature 8. Notifications inform users of new actions such as a play, pause, or subpage change (Feature 7). Features 5 and 9 show the webcam feed of all users. Friends can be invited into the living room by clicking the invite-button to the left of participating users. Feature 10 is the text chat, where both users have written a message. All users of the room will always view the same of these two screens containing the same content at any point in time.

In early iterations of the system, all users could control the selection and playback of videos simultaneously. However, test sessions showed that this seemed to cause confusion and chaotic experiences. In order to better facilitate control over the experience, we created a virtual "remote control". The remote control is assigned only to one user, who is then able to browse and search for video content and interact with the video player to start and stop playback of videos (see feature 5 in Figure 1 and feature 9 in Figure 2). If the user wants to hand over control to one of the other users, they can"pass" the remote control to the other user simply by clicking on it. While this feature arguably creates an artificial limitation on the users' interactions, user testing showed that it greatly improved coordination among people, and they quickly picked up on the metaphor and learned to pass the remote control around.

While much of the design process was aimed at emulating the experience of watching TV together as one would do in a physical living room, we also explored ideas around how an online experience could offer *new* social interactions with video content. As a result, we created a drawing feature that was intended to encourage discussion by allowing users to pause the video and draw directly on the screen using their pointer device (see Figure 3). The idea was based on feedback from users in early iterations, who expressed a wish to be able to pause and point at a specific spot in the video, analogous to pointing the finger at the television in the physical living room. This feature is only activated when the video is paused. Users draw by using their pointer (e.g. the mouse or their finger) on the video screen, and their strokes immediately appear on the screen of every user in the virtual living room.

An issue of great importance for this system is synchronization of the video content. We believe it is important that the content is synchronized with great precision, in order to facilitate an experience of togetherness. If one person reacts to an event on screen

slightly before the other - such as cheering for a goal in a soccer match, or laughing at a joke in a comedy - this would be likely to disrupt their feeling of a shared experience. To avoid this problem, the system uses a peer-to-peer connection to synchronize both the webcam/voice feeds as well as the video content, so that any delay in the video content will be matched by an identical delay in the webcam/audio connection. The actual delay depends on the users' internet connection, but most users in our country have only 1-30 ms delay. No test users experienced synchronization problems.

The final concept, as described above, and illustrated in figures 1 and 2, was released as a publicly available system for anyone to use in April 2020. As opposed to previous video sharing systems, ours was readily available on the web, required no app installation and worked with both YouTube and the national broadcaster's web player. At this point the system was trialled in a more detailed fashion, using the methods described above.

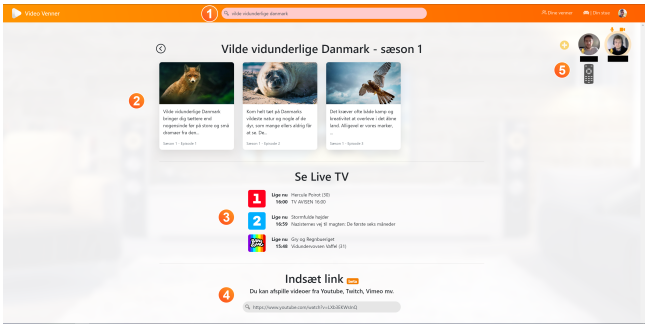

Figure 1: The video selection screen of the virtual living room. Content may be selected by searching, by clicking a live television channel or by a video link to either Facebook or Youtube. (The names of the two users have been redacted.)

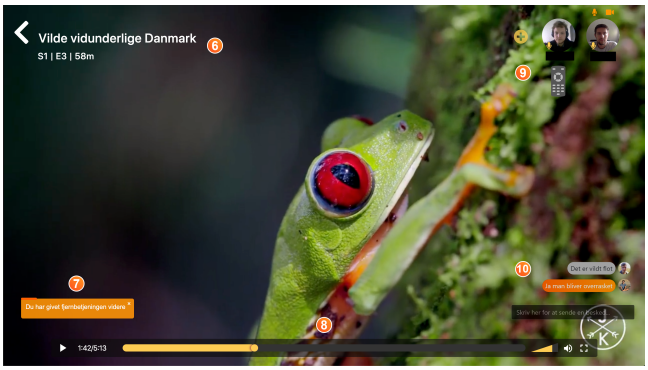

Figure 2: The video player screen of the virtual living room where all participants of a room watch the same video synchronized. Note that the feature count is a continuation from Figure 1. (Parts of the image has been blurred for anonymity.)

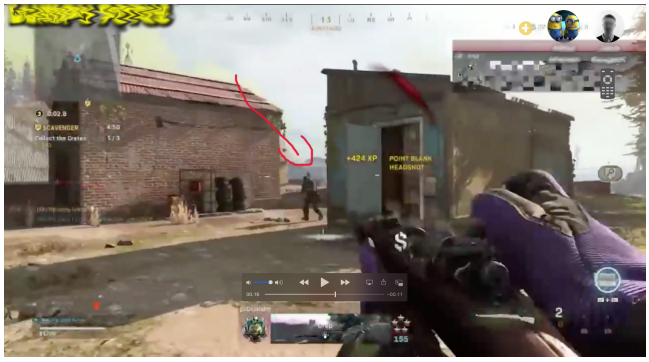

Figure 3: The drawing feature. The user has paused the video and drawn a red arrow to indicate a particular detail to the other user, while discussing the video. (The names of the two users have been redacted.)

5 EVALUATION

In the following we present our evaluation methodology - first, regarding the formative and iterative user testing throughout the design process, second regarding the summative evaluation at the end of the design process. Subsequently we present the main insights from the evaluation of the final prototype.

5.1 Iterative User Tests

We conducted a total of 20 user tests, throughout the development, with 35 different participants. Each user test included between two and four participants, mostly two. In five of the tests, participants interacted with one of the authors. Participants were set up with the prototype system in two different rooms (in a few later iterations fully remote) and asked to "watch something together", meaning they got to chose what they wanted to watch themselves. We used observation and user logs for getting user data. These user tests were useful for fine-tuning design, especially the user interface and for getting feedback on possible elements of the system. Although we have documented the detailed findings of these, we focus in this paper on the larger final evaluation that provided insights into broader issues around sociability and video content.

5.2 User Evaluation Strategy

We evaluated the system with users over a period of four weeks in April-May 2020. While the COVID-19 pandemic made many people unable to socialize with their friends and family, this provided a good opportunity to launch the prototype to the public and obtain user feedback from use "in the wild". We envisioned that small groups of friends and family would find the prototype useful and in order to gather insights into their user practices and opinions of the prototype, we distributed a survey. This was implemented as a popup that was added to the prototype, which was automatically displayed when users of a virtual living room would leave after watching a video, asking them to participate in the survey. Over a period of four weeks, a total of 131 users registered a profile, and at least 40 users visited the virtual living room with at least one friend. 11 users replied to the survey after visiting the virtual living room.

In order to obtain more in-depth insights about the experience, the 11 respondents were contacted using the e-mail addresses they had provided, and asked if they wanted to participate in an interview and observation session with their friend or friends that they were watching with. From this, two observation sessions with a total of six users were conducted through video link. Furthermore, three observations were conducted with users recruited through the authors' own networks (snowball sampling). For the user testing, it was arranged such that one of the test users would share their screen through a remote video link (Skype) with the researchers who would remain on mute through the entire observation session to mimic physical fly-on-the-wall observations. Participants chose themselves which video content to watch, among the possibilities offered by the system. For privacy reasons both webcam and voice chat were disabled as the users started their sessions, however all users chose to enable webcam and voice. Sessions lasted 30-50 minutes. After the user session, the participants were interviewed through the same channel, which was recorded; interviews lasted 20-30 minutes. Participants each received a gift card worth the equivalent of USD 15. Observational test and interviews were conducted on a total of five pairs between the ages 24 and 61, none of which had participated in earlier user tests (the participants' names have been replaced with pseudonyms):

- Anna and Betty
- Charlie and Carla, David and Donna (two married couples)
- Eric and Frank
- Gina and Holly
- Irina and Julie

Before presenting the results we describe in more detail the design process and final prototype of the video sharing system.

5.3 Interacting with the System

All users in the five interactive video sessions were observed interacting with the system without major usability issues. For example, without any prior introduction to the concept of the digital remote control, they were observed to "pass" the remote control to each other without needing instructions. Irina and Julie said that the remote control was a nice feature as it encouraged more communication between users, and that it clearly indicated who was in control, analogous to the physical living room. Eric and Frank had similar positive experiences with the remote, with Frank for example saying: "It was a bit easier when only one could search; otherwise, it would have been rather confusing. It gives better control and peace of mind" (Frank).

The observations also showed that the majority of the users preferred webcam communication over just speech, and all users preferred speech and webcam over textual communication. Text communication was only used to coordinate prior to entering into a digital living room, and as a fallback when technical issues occurred. These observations resonate with similar findings presented by Macaranas et al. [22]. One smaller issue in relation to the design that emerged through the interviews was a suggestion for increasing the webcam viewport size: "[...] I would just like to be able to make the picture of us bigger" (Julie).

We were very curious how the users would approach the drawing feature, particularly since this was the most novel aspect of our video sharing prototype. The drawing feature was used in three of the five sessions we observed, and each time it helped facilitate conversation about the video the users were watching. We observed how the participants quickly adopted the drawing feature in an intuitive way, for example using the feature to describe or discuss a particular event in a video. Figure 3 shows Eric and Frank watching a video of a computer game session that Frank had already watched but wanted to share with Eric. Frank had paused the video and used the drawing feature to point out an enemy that had just appeared from behind a corner, exclaiming: "Look, and now another person appears, just there!" Eric responded: "He's just got two bullets left, no way he'll get him." When Frank starts the video and they see that the player kills the opponent with a single well-placed shot, they both laugh and cheer in amazement. Later in the same session, Frank pauses the video to draw a circle in an apparently empty area of the video game, telling Eric: "Look at the reaction time of this guy. Another player is going to come here [highlighted area], and he just reacts so fast". The speed and ease with which these users integrated the drawing feature into their dialogue about the video indicates that such feature is a very promising addition to remote video watching.

5.4 Experiences of the Social Interaction

One of the issues we wanted to look into was the organic experience of socializing (around television) through online media. Similarly to Ducheneaut et al. [10], we observed that users' interactions were tightly interwoven with the structure of the video they were watching - for instance, they timed their communication to fit within pauses in dialogue and transitions. When inquiring into the social experience, however, opinions varied. Eric gave a positive view, indicating that the experience compared well to watching physically together: "I think it is actually a lot easier and faster to see video together online; to just hop into a room and put on a show, and then you can always just stop if it gets boring. It is very noncommittal" (Eric).

However, this was not a general perception among all the people who had tried the system. In the survey, users were asked to rate their shared remote video experience compared to a co-located viewing experience, on a 5-point scale from "very different" to "very similar" and responses varied widely: Of the 11 respondents, 4 answered "somewhat different" or "very different", whereas 3 answered "somewhat similar" (and no-one said "very similar"). However, when users were asked to rate how similar the experience was to watching alone, 8 respondents rated it as "somewhat different" or "very different", and only 1 answered "somewhat similar" (while no-one responded "very similar"). Finally, when asked to rate how well the communication worked during their shared remote viewing, 10 of the 11 respondents answered "well" or "very well", and no-one answered "poorly" or "very poorly". These survey data indicate that users rate the experience to be quite different both from watching alone and from watching together in the physical living room, suggesting that watching video together online might be categorized as a third experience separate both from watching alone and physically together.

In early iterations, prior to implementing the virtual remote control, some users also described the experience as being more

chaotic than in the physical living room. This might be considered analogous to having a physically co-located TV night in which every person has their own remote-control, providing everybody with the ability to decide on what and when to watch. While this may encourage participation, it could also easily result in more disorganized experiences. The virtual remote control we implemented to address this problem seemed to have helped ground the experience in the users' existing living room practices, and helped facilitate a less chaotic experience. For instance, one participant said: "I think it worked really well [the digital remote control]. It makes sense considering it is supposed to be a living room, and it makes sure that anyone can't just change the video" (Eric). A few users, however, argued that it would fit the living room analogy better if users could simply take the remote: "I think it is good that we both can't pause it at the same time. But if Anne had to go to the toilet, then she would have to ask for permission. In the real world, you can just grab the remote and pause" (Betty).

A particularly interesting reflection came from one user, who pointed out that this system may facilitate both shorter, casual sessions as well as longer "movie night" experiences:

> "It is strange because you are sitting at home in your empty room but still feel some kind of closeness to the person you are watching together with. You don't feel that you are alone. Even though it is not like a movie night, I often share some funny videos on Instagram or Facebook with Anne, and now we can watch these together, for example, if we don't want to watch a whole episode of something." (Betty)

This description is interesting because it illustrates a different use-case of the system than the "movie night" experience originally envisioned by the designers: Instead of getting together for a long session to watch a movie or a tv show, the user suggests she would want to use the system to quickly connect with a friend to watch a short video clip from social media together. Eric expressed a similar statement: "If I am with someone here in my own living room, it is about the close connections. On the other hand, here [in the virtual living room] I can simply quickly share something with a friend" (Eric). Another user, David, described a scenario in which the video content created a common ground for conversation and made their online social experience less artificial: "It has been great to be able to be with Charlie and [his wife] and watch a movie. It quickly becomes very artificial if we are on Skype without doing something simultaneously" (David).

Perhaps unsurprisingly, when the users were asked if the system could replace watching videos together physically they tended to say no, but rather saw it as a substitute when meeting physically is not an option - such as during the current COVID-19 pandemic. In fact, one user hosted a family dinner during the Easter holiday using our prototype:

> "We ate dinner together each to ourselves on here, and then we could just put on a movie afterward. We had dinner with the camera on. And then we did not even have to drive home after the movie, so it was really good we could sit each to ourselves, especially with Corona around." (David)

This suggests that users may want to use a remote video watching system such as LetsWatch for a broad variety of experiences, ranging from casual sharing to longer-lasting "full night" experiences.

6 COMMERCIAL DEPLOYMENT

Since the end of our research project, two of the authors have been working with the national broadcaster to further develop the prototype and implement it as a feature in the broadcaster's online video player. The feature went through extensive testing and a beta trial, before being launched to the general public in September 2021. Within the first month of deployment the number of users surpassed the benchmark set by the broadcaster, and representatives of the broadcaster have deemed the project a success. Initially the software was offered free of charge to the broadcaster on a trial basis, but it has since become a paid service. The system runs independently of the broadcaster's systems and may also be deployed elsewhere - negotiations are underway with several other broadcasters.

Due to the broadcaster's requirements that the implemented feature must pass through rigorous user testing, the implementation initially left out some of the functionality of the prototype described earlier in this paper. Notably, communication between users has initially been restricted to text chat.

Below we present some usage statistics covering the time period from 7 October 2021 to 1 February 2022. These offer both some early indication that users are finding the feature useful, as well as some insights into the ways users are adopting the feature.[9]

- Number of users: 60231
- Number of unique users: 20650
- Number of VOD Rooms: 23016
- Number of Live TV Rooms: 1891
- Average duration of room session: 39 minutes
- Average length of message (in characters): 12
- Average number of messages per active room: 40
- Average number of users per active room: 2.3

As can be seen from these numbers, the feature is indeed used as a social function to connect users - mostly 2 users watching together but also a substantial number of sessions with 3 or more users, leading to an average of 2.3 users per room. While the number of users demonstrate that for now this is a niche feature used by a relatively small group of users, the numbers also seem to point towards extensive engagement from these users, with fairly long average duration of sessions (39 minutes), and a nearly 3:1 difference between total users and unique users indicating that many users return to the feature repeatedly. The text chat is also used extensively with 40 messages per room on average - and the average message length of 12 characters may indicate that the conversation goes beyond just short exclamations and emojis. Somewhat surprising to us is the large domination of VOD (Video-on-Demand) content over live TV, with a ratio of 12:1 in the number of VOD rooms versus live rooms. It should be noted that on some occasions this balance has shifted, such as during the UEFA European Football Championship in the summer of 2021, as well as during some of the prime minister's press conferences regarding the COVID-19

[9]Note that there is no login required to use the LetsWatch feature, which places some limitations on the data that can be acquired.

pandemic. Both of these examples represent large media events which were seen live by large audiences - this was reflected also in usage numbers for the LetsWatch feature. However, in the time period covered by the statistics above, the most watched content were the daily "Christmas Calendar" shows during 1-24 December 2021. These shows are primarily directed at children and families.

Attracting young audiences was one of the broadcaster's main motivations for implementing the LetsWatch feature. Since the feature does not require a login we cannot measure the age distribution of users directly. However, we can get some indication by looking at the answers to a feedback survey which is presented to users when they leave a room. The numbers below show the proportion of responses from different age groups, separated in standard age brackets used in the broadcaster's audience research (N=116):

- 0-8 years: 14%
- 9-14 years: 19%
- 15-24 years: 27%
- 25-31 years: 16%
- 32-46 years: 11%
- 47-64 years: 13%
- 65+ years: 1%

About 60% of the responses come from users below the age of 25, indicating a great dominance of young viewers. If this reflects the age distribution among all the users this is a very successful result for the broadcaster, which is struggling to capture audiences among teens and young adults.

When users were asked who they watch together with, most refer to either friends, family or romantic partners, confirming that this feature is used primarily to share an experience with their strong social ties. As shown in Figure 4, use of the feature at different times of a typical day corresponds roughly to typical TV audience numbers, with moderate activity during the day climbing after work hours to a peak between 7pm and 10pm in the evening.

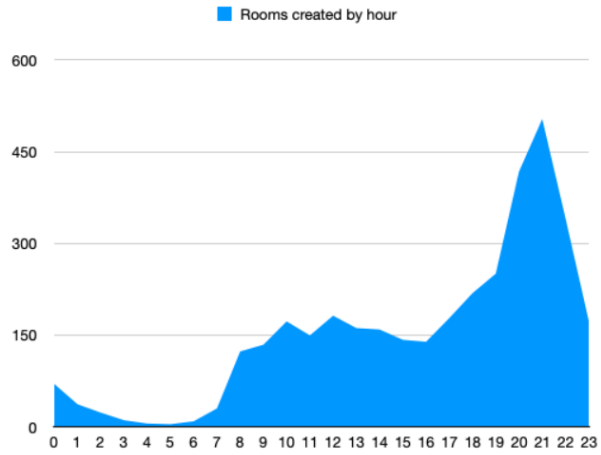

Figure 4: Usage of the implemented feature on the national broadcaster's video player on a typical day. The graph is normalised so that the y-axis does not represent actual number of rooms created, but rather just the distribution of activity during the day.

7 DISCUSSION

The video sharing system presented in this paper differs from the many research prototypes that have been presented in past research (see Section 2 on Social Watching Built for Research above), in that the LetsWatch system has been deployed to a mass audience and thereby offers some proof that the concept is deemed relevant and valuable for the TV industry, as well as offering "in-the-wild" data from everyday use. Viewed in this perspective, it is interesting to note that the average session of 39 minutes seems to indicate quite long-lasting engagement, in a situation where people may be entering the service just to try it out. While this is encouraging, there also seems room for improvement and further research. Here we highlight in particular issues relating to the control of content, as well as the drawing feature.

Arguably, using the metaphor of a physical remote control to facilitate coordination of remote watching replicates a limitation from the physical world which might be unnecessary in an online format. While this design is in line with recommendations in past research [23, 29], alternative solutions that allow users to share control simultaneously have been proposed in research [31] and have also been implemented in commercial services such as the Watch2Gether system (see section on State of the Art). Our experiences from the design process and iterative user testing reflect the findings of McGill et al. [23] who conclude that systems which offer control to one user at a time give better usability. However, future research might challenge the trade-off between improving usability and restricting control, in order to explore more flexible solutions that do not replicate restrictions from the traditional TV apparatus. For instance, designers might implement mechanisms for passing the control around through voting or turn-taking; systems might also allow users to vote for the next video, or even turn the selection and control of content into playful or game-like experiences.

The drawing feature in LetsWatch could also be seen as replicating an interaction that might easily take place in the physical living room - pausing the video and pointing at something on the screen. This is an interaction that may be particularly relevant for "casual" video watching practices that have become increasingly popular in recent years, such as watching videos of computer games, or watching short videos from social media and video sharing websites such as YouTube. Furthermore, this feature might be useful for educational or instructional settings - e.g. the prototype could be used for screening educational material to a small group of students, and the teacher might use the drawing feature to point out important details. The system might be a useful platform for educational settings, allowing school children to experience the co-presence of their classmates when forced to carry out school activities from home (such as during the current pandemic, or due to other illness). Thus the drawing feature may be used as a tool that turns the mostly passive TV watching experience into a more active and participatory experience. However, the drawing feature raises a number of questions for further design research:

- Is there a need to manage control of this feature? Currently, once the video is paused anyone can draw on the screen, and

the drawing instantly becomes visible to all other users in the room. This is not likely to be a problem in the types of settings we have explored here, where each room consists of few users who know each other well - but might be a problem in other settings, e.g. with larger and more mixed groups (e.g. if used in educational settings).
- Should it be possible to use the drawing feature while the video is playing? One might imagine situations where this could be relevant, e.g. for instructional use, or in relation to sports or video games. In such cases it might also be desirable to be able to point at something without creating a drawing, similar to a "laser pointer" used in classrooms.
- In the current version, once the video starts playing again the drawing is erased. If users could save their drawings the feature could also be used as a tool for making visual commentary (or memes) based on video content, that could be downloaded and shared. Future research might explore how to develop this further as a collaborative creative tool.

The LetsWatch system currently mainly facilitates a shared video experience for users who already know each other - it has no social networking features that might help users meet new people. As a theme for further development, one might explore further ways of connecting different users. For instance, one might set up "screening rooms" where people could watch movies or other content together with strangers - mimicking offline practices such as film clubs, sports pubs, etc. However, such work falls outside the scope of this project.

8 CONCLUSION

In this paper we described our experiences of designing and testing, in real settings, a social video sharing system called LetsWatch. Our prototype demonstrated relevant insights about the specific features that may be deployed to facilitate a shared video experience. In particular, the virtual remote control was shown to work as a well-known metaphor for coordinating control of the video selection and playback. Furthermore, the drawing feature was shown to be a feature that facilitated dialogue and allowed more interactive conversations about video content.

Our study also shed some light on how the system facilitated the users' perception of such a remote shared video experience. In general, the experience of the virtual living room was found to be different, both from the experience of watching alone and from that of co-located viewing parties, instead described as a *third* experience and a valuable supplement, especially when meeting physically is not an option. The experience was found to add a sought after social aspect compared to watching alone, and to establish a feeling of co-presence sufficient to facilitate strong video-mediated sociability. Furthermore, the virtual living room was found to provide new use cases compared to the physical living room, by additionally enabling brief and casual viewing sessions.

This study occurred at a moment in time when the COVID-19 pandemic was forcing large populations of people across the world to engage in social distancing, avoiding co-present socialising with most of their networks. While this context has demonstrated an acute need for new technologies for remote socialising, this need is likely to remain relevant even after the pandemic.

The implementation of the system in the website of the national broadcaster has demonstrated that such a system is interesting for broadcasters, in particular due to its potential to engage young people - a potential that appears to have been confirmed in our data. But the need to connect with others is great also among older age groups. An important challenge for future work will be to explore how to make such a system appeal broadly across age groups, allowing for richer distributed socializing experiences with video content.

ACKNOWLEDGMENTS

We thank all the test users who helped in evaluating the design, as well as the anonymous reviewers for their thoughtful contributions.

REFERENCES

[1] Morgan G. Ames, Janet Go, Joseph 'Jofish' Kaye, and Mirjana Spasojevic. 2010. Making Love in the Network Closet: The Benefits and Work of Family Videochat. In *Proceedings of the 2010 ACM Conference on Computer Supported Cooperative Work* (Savannah, Georgia, USA) *(CSCW '10)*. Association for Computing Machinery, New York, NY, USA, 145–154. https://doi.org/10.1145/1718918.1718946

[2] Lynne Baillie, Peter Frohlich, and Raimund Schatz. 2007. Exploring Social TV. In *2007 29th International Conference on Information Technology Interfaces*. 215–220. https://doi.org/10.1109/ITI.2007.4283773

[3] Louise Barkhuus. 2009. Television on the Internet: New Practices, New Viewers. *Association for Computing Machinery, CHI '09 Extended Abstracts on Human Factors in Computing Systems* (2009).

[4] Louise Barkhuus and Barry Brown. 2009. Unpacking the Television: User Practices around a Changing Technology. *ACM Trans. Comput.-Hum. Interact.* 16, 3, Article 15 (Sept. 2009), 22 pages. https://doi.org/10.1145/1592440.1592444

[5] Louise Barkhuus, Arvid Engström, and Goranka Zoric. 2014. Watching the footwork: Second screen interaction at a dance and music performance. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 1305–1314.

[6] Margot Brereton, Alessandro Soro, Kate Vaisutis, and Paul Roe. 2015. The Messaging Kettle: Prototyping Connection over a Distance between Adult Children and Older Parents. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (Seoul, Republic of Korea) *(CHI '15)*. Association for Computing Machinery, New York, NY, USA, 713–716. https://doi.org/10.1145/2702123.2702462

[7] Pablo Cesar and David Geerts. 2011. Understanding Social TV: a survey. *Proceedings of the Networked and Electronic Media Summit (NEM Summit 2011), Torino, Italy, September 27-29* (2011).

[8] The Nielsen Company. 2021. The Gauge Shows Streaming is Taking a Seat at the Table. (2021). https://www.nielsen.com/us/en/insights/article/2021/the-gauge-shows-streaming-takes-a-seat-at-the-table/#methodology

[9] Anind K. Dey and Ed de Guzman. 2006. From Awareness to Connectedness: The Design and Deployment of Presence Displays. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Montréal, Québec, Canada) *(CHI '06)*. Association for Computing Machinery, New York, NY, USA, 899–908. https://doi.org/10.1145/1124772.1124905

[10] Nicolas Ducheneaut, Robert J. Moore1, Lora Oehlberg, James D. Thornton, and Eric Nickell. 2008. Social TV: Designing for Distributed, Sociable Television Viewing. *Palo Alto Research Center and UC Berkeley, Mechanical Engineering* (2008).

[11] Tom Feltwell, Gavin Wood, Scarlett Rowland, Kiel S. Long, Chris Elsden, Phillip Brooker, John Vines, Pamela Briggs, Julie Barnett, and Shaun Lawson. 2019. Designing Second-Screening Experiences for Social Co-Selection and Critical Co-Viewing of Reality TV. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3290605.3300300

[12] David Gauntlett and Annette Hill. 1999. *TV Living: Television, Culture and Everyday Life*. Routledge; 1 edition (June 23, 1999).

[13] David Geerts, Ishan Vaishnavi, Rufael Mekuria, Oskar van Deventer, and Pablo Cesar. 2011. Are We in Sync? Synchronization Requirements for Watching Online Video Together. *Proceedings of the International Conference on Human Factors in Computing Systems, CHI 2011, Vancouver, BC, Canada, May 7-12, 2011* (2011).

[14] Fenghua Guo, Caiming Zhang, and Lizhen Cui. 2007. Sketching Interfaces for Remote Collaboration. In *2007 11th International Conference on Computer Supported Cooperative Work in Design*. 63–68. https://doi.org/10.1109/CSCWD.2007.4281411

[15] Gunnar Harboe, Crysta J. Metcalf, Frank Bentley, Joe Tullio, Noel Massey, and Guy Romano. 2008. Ambient Social Tv: Drawing People into a Shared Experience. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*

(Florence, Italy) *(CHI '08)*. Association for Computing Machinery, New York, NY, USA, 1–10. https://doi.org/10.1145/1357054.1357056

[16] Jim Hollan and Scott Stornetta. 1992. Beyond Being There. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Monterey, California, USA) *(CHI '92)*. Association for Computing Machinery, New York, NY, USA, 119–125. https://doi.org/10.1145/142750.142769

[17] Jake Knapp, John Zeratsky, and Braden Kowitz. 2016. *Sprint: How to Solve Big Problems and Test New Ideas in Just Five Days*. Transworld Publishers Ltd.

[18] Peter Gall Krogh and Marianne Graves Petersen. 2010. *Designing for Collective Interaction: Toward Desirable Spaces in Homes and Libraries*. Springer London, London, 97–113. https://doi.org/10.1007/978-1-84882-965-7_5

[19] Hong Li, Heiko Müller, and Jonna Häkkilä. 2020. Our Little Secret: Design and User Study on an Electrochromic Ambient Display for Supporting Long-Distance Relationships. In *Interactivity, Game Creation, Design, Learning, and Innovation*, Anthony Brooks and Eva Irene Brooks (Eds.). Springer International Publishing, Cham, 611–622.

[20] Yiming Liu, Peter Shafton, and Jeannie Yang. 2008. Zync: the design of synchronized video sharing. *International Journal of Human–Computer Interaction, 24:2, 136-154* (2008).

[21] James Lull. 1990. *Inside Family Viewing: Ethnographic Research on Television's Audiences*. London: Routledge.

[22] Anna Macaranas, Gina Venolia, Kori Inkpen, and John Tang. 2013. Sharing Experiences over Video: watching video programs together at a distance. *Human-Computer Interaction – INTERACT 2013 : 14th IFIP TC 13 International Conference, Cape Town, South Africa, September 2-6, 2013* (2013).

[23] Mark McGill, John Williamson, and Stephen A. Brewster. 2014. How to Lose Friends amp; Alienate People: Sharing Control of a Single-User TV System. In *Proceedings of the ACM International Conference on Interactive Experiences for TV and Online Video* (Newcastle Upon Tyne, United Kingdom) *(TVX '14)*. Association for Computing Machinery, New York, NY, USA, 147–154. https://doi.org/10.1145/2602299.2602318

[24] Mark McGill, John H. Williamson, and Stephen Brewster. 2016. Examining The Role of Smart TVs and VR HMDs in Synchronous At-a-Distance Media Consumption. *ACM Trans. Comput.-Hum. Interact.* 23, 5, Article 33 (Nov. 2016), 57 pages. https://doi.org/10.1145/2983530

[25] Ian Morris. 2017. Netflix Is Now Bigger Than Cable TV. *Forbes* (June 2017). https://www.forbes.com/sites/ianmorris/2017/06/13/netflix-is-now-bigger-than-cable-tv/ Section: Tech.

[26] Mukesh Nathan, Chris Harrison, Svetlana Yarosh, Loren Terveen, Larry Stead, and Brian Amento. 2008. CollaboraTV: Making Television Viewing Social Again. In *Proceedings of the 1st International Conference on Designing Interactive User Experiences for TV and Video* (Silicon Valley, California, USA) *(UXTV '08)*. Association for Computing Machinery, New York, NY, USA, 85–94. https://doi.org/10.1145/1453805.1453824

[27] Jiazhi Ou, Xilin Chen, Susan R. Fussell, and Jie Yang. 2003. DOVE: Drawing over Video Environment. In *Proceedings of the Eleventh ACM International Conference on Multimedia* (Berkeley, CA, USA) *(MULTIMEDIA '03)*. Association for Computing Machinery, New York, NY, USA, 100–101. https://doi.org/10.1145/957013.957034

[28] Hayes Raffle, Glenda Revelle, Koichi Mori, Rafael Ballagas, Kyle Buza, Hiroshi Horii, Joseph Kaye, Kristin Cook, Natalie Freed, Janet Go, and Mirjana Spasojevic. 2011. Hello, is Grandma There? Let's Read! StoryVisit: Family Video Chat and Connected e-Books. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Vancouver, BC, Canada) *(CHI '11)*. Association for Computing Machinery, New York, NY, USA, 1195–1204. https://doi.org/10.1145/1978942.1979121

[29] David A. Shamma, Marcello Bastea-Forte, Niels Joubert, and Yiming Liu. 2008. Enhancing Online Personal Connections through the Synchronized Sharing of Online Video. In *CHI '08 Extended Abstracts on Human Factors in Computing Systems* (Florence, Italy) *(CHI EA '08)*. Association for Computing Machinery, New York, NY, USA, 2931–2936. https://doi.org/10.1145/1358628.1358786

[30] Vikash Singh, Celine Latulipe, Erin Carroll, and Danielle Lottridge. 2011. The choreographer's notebook: a video annotation system for dancers and choreographers. In *Proceedings of the 8th ACM Conference on Creativity and Cognition*. 197–206.

[31] Emily Sun, Rodrigo de Oliveira, and Joshua Lewandowski. 2017. Challenges on the Journey to Co-Watching YouTube. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing* (Portland, Oregon, USA) *(CSCW '17)*. Association for Computing Machinery, New York, NY, USA, 783–793. https://doi.org/10.1145/2998181.2998228

[32] Maarten Wijnants, Jeroen Leën, Peter Quax, and Wim Lamotte. 2014. *Augmented Video Viewing: Transforming Video Consumption into an Active Experience*. Association for Computing Machinery, New York, NY, USA, 164–167. https://doi.org/10.1145/2557642.2579368

[33] Doug Williams, Marian F Ursu, Pablo Cesar, Karl Bergström, Ian Kegel, and Joshan Meenowa. 2009. An emergent role for TV in social communication. In *Proceedings of the 7th European Conference on Interactive TV and Video*. 19–28.