DESIGNING INTELLIGENT INTERFACES TO FACILITATE COMMUNICATION IN LARGE-SCALE LIVE STREAMING

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Jiajing Guo

May 2022

DESIGNING INTELLIGENT INTERFACES TO FACILITATE COMMUNICATION IN LARGE-SCALE LIVE STREAMING

Jiajing Guo, Ph. D.

Cornell University 2022

Live streaming is a type of medium that integrates video, text chat, and even more modalities to support real-time communication. It has become increasingly popular and accessible in recent years. Large-scale live streams face challenges in both real-time text chat and friendship development in the community. My dissertation explores the design space of intelligent interfaces to facilitate communication in large-scale livestreams.

I conducted a qualitative interview study that investigated the motivation, practices, and challenges people face when engaging with the streamer and peer viewers on social network sites. Based on the findings, I applied the similarity-attraction effect from social psychology and proposed a subgrouping method that generates subgroups for viewers based on their similarities. Then I designed and implemented Chatbuddies, an intelligent interface that generates subgroups and interactively visualizes viewers' information.

In a controlled within-subject lab experiment and an unmoderated online experiment, participants reported that chats in subgroups were easier to follow and more enjoyable than chats in a channel without subgroups. They perceived peer viewers as more attractive in a similarity-based subgroup than in random groups. This dissertation contributes to the understanding of challenges in large-scale communication and provides design implications for future AI-embedded CMC.

BIOGRAPHICAL SKETCH

Jiajing Guo completed her Ph.D. in the Department of Information Science at Cornell University, with a minor in Computer Science in 2022. She worked in the Communication and Collaborative Technologies Lab led by Prof. Susan R. Fussell. Prior to Cornell, she received a bachelor's degree in Information Design at Tsinghua University, Beijing in 2017. Jiajing was originally from Taizhou, a tranquil and beautiful city in Jiangsu Province, China. In addition to research, she loves to practice yoga and create illustrations.

ACKNOWLEDGMENTS

I want to thank the amazing people who have helped and supported me in the amazing journey over the past five years. First and foremost, I am forever grateful to my advisor, Dr. Sue Fussell, who has been a brilliant mentor and supporter throughout my doctoral research. This dissertation could not have been possible without her enormous patience to grow me as an independent researcher and huge engagement when I was confused and lost. Thinking of the days and nights where we chatted about research projects, experiment stats, paper edits, and everything else, I would like to cherish it as one of my most precious memories in my life.

I would also like to thank my committee, Dr. Malte Jung, and Dr. Claire Cardie for their insightful and detailed comments on my research. I enjoyed the walking meetings with Malte and his challenges that encouraged me to expand my thinking and design. I am honored to be in Claire's research group, to learn up-to-date NLP knowledge, and to collaborate with amazing undergrads.

Beyond my committee members, I would like to thank faculty members who I have worked with as a teaching assistant, Dr. Gilly Leshed, Dr. François Guimbretière, Dr. Kyle Harms, Dr. Cheng Zhang, and Dr. Jeff Rzeszotarski who showed me enthusiasm in teaching and research. I also want to thank the amazing administrative staff in the Information Science Department, Barbara Ann Warner and Terry Horgan for all of their help with paperwork and the logistics of being a graduate student.

I want to thank my lab mates in the Communication and the Collaborative Technologies Lab who provided me with invaluable feedback and spent time discussing research ideas and technical details. They are Wen Duan, Alex Hinck, Negar Khojasteh, Jingjin Li, Xiaoyan Li, Erica Ostermann, Sharifa Sultana, Neta Tamir, Luping Wang, and Elijah Weber-Han. Beyond the lab, I want to thank fellow graduate students in my department and friends I made in Ithaca for their invaluable instrumental and mental support. Further, I want to thank all my research assistants for their dedication and hard work. They helped me reflect on my own research and learn to be a good mentor. Without their efforts, my Ph.D. research will not be successful.

Last but not least, I would like to thank all of my family members who always respected my choice and supported me no matter what happened. My parents are my role models of having a positive mind and being dedicated to things they love. My father spent countless hours accompanying me when I practiced calligraphy in childhood, which shaped my calm and strong mind. My mother is a lifelong friend, who always gives me sincere suggestions and guides me to see things from another side. I also want to thank my grandparents and my aunts' family for their blessings and love. Finally, a special thank goes to my partner, Kai Yao, who has been a source of constant support since high school. He reminded me the importance of enjoying life, shared my happiness and grief, and gave me tremendous support during my dissertation writing.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

AI	Artificial Intelligence
CMC	Computer-mediated Communication
CSCW	Computer-supported Cooperative Work
HCI	Human-computer Interaction
ML	Machine Learning
NLP	Natural Language Processing
SNS	Social Network Sites

CHAPTER 1

INTRODUCTION

The advent of Internet and computer-mediated communication (CMC) technologies revolutionized interpersonal communication by providing individuals with a variety of channels and modalities to interact across time and regions (Herring, 2002). Before the emergence of CMC, real-time communication was constrained by physical space, therefore face-to-face interaction generally happened in dyads or small groups (Whittaker et al., 2003), except in some scenarios such as lectures and public speaking. Explosive growth in CMC tools shrank distances and facilitated information exchange among people in various locations by creating virtual space and supporting not only one-to-one but large-scale interaction with hundreds of users involved at the same time.

Nowadays such large-scale events are often held in the form of live streaming, a type of media where live video and text chat are integrated to support real-time communication. Figure 1 is a screenshot of a Twitch live stream of a video game. Figure 2 is a screenshot of a YouTube live stream One World: Together at Home concert, broadcast on April 18, 2021. A typical live stream interface on Desktop devices is split into two components. The left side is usually a high-fidelity live broadcasting video where streamers share gameplay or real life. The right side is a text chat window that supports real-time communication for viewers. Streamers may engage with viewers via video and audio. Some platforms have advanced features such as inviting guests (J. Li et al., 2019), gifting and donation (Lee et al., 2018; Lu, Xia, et al., 2018).

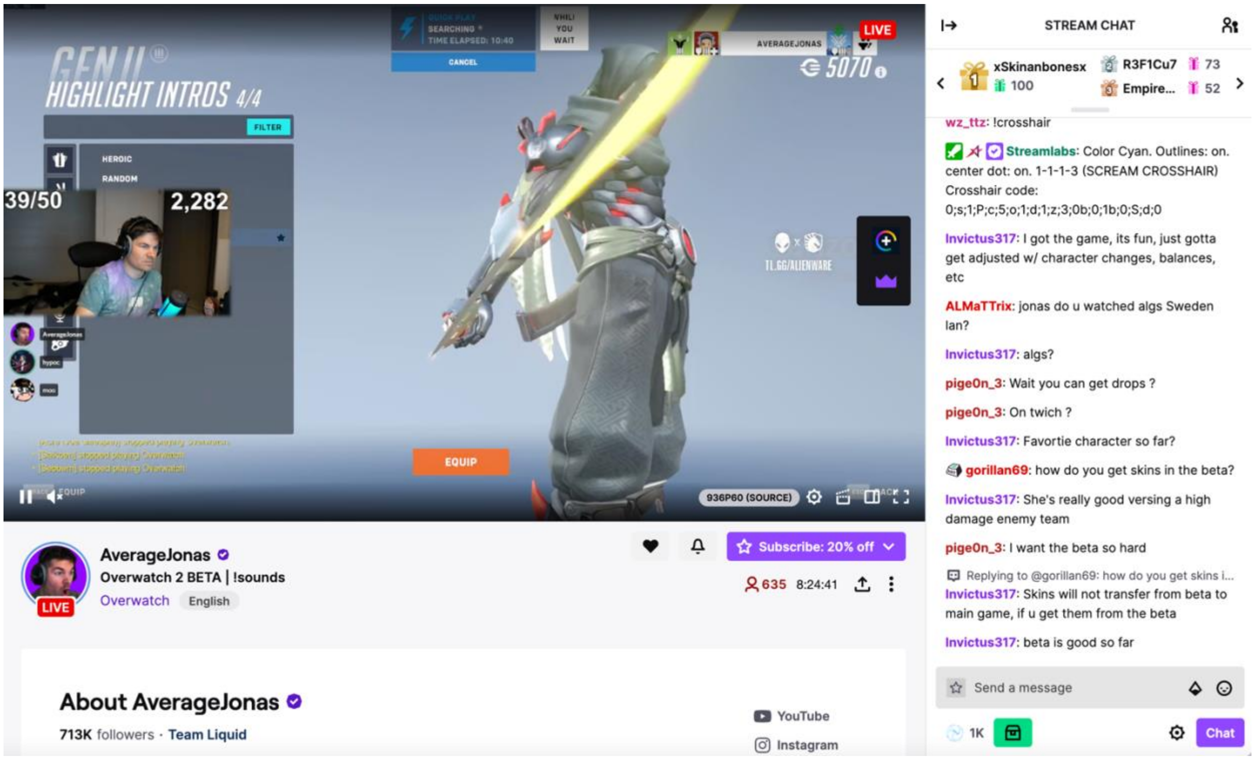


Figure 1: A Twitch live stream has a live video on the left and chats on the right.

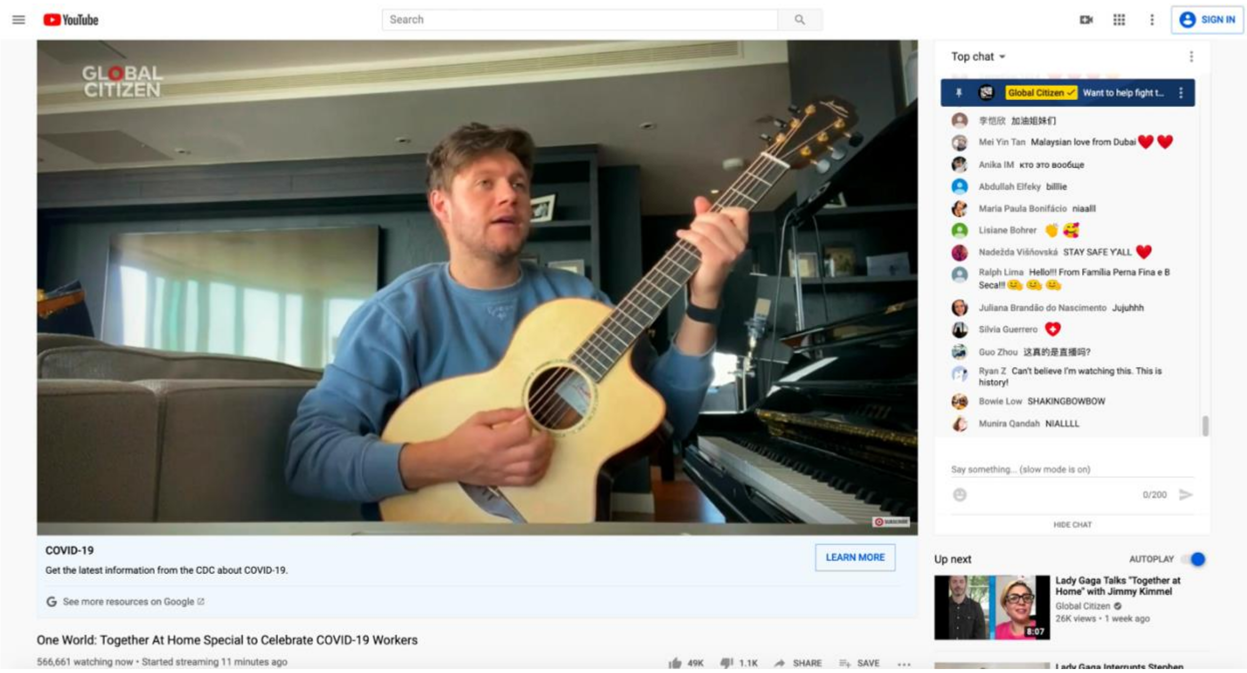


Figure 2: A YouTube live stream has a live video on the left and chats on the right

While live stream is not a new concept (Tang et al., 2016), it became increasingly popular in recent years as it was embedded into many popular social network sites, such as Facebook (Meta, n.d.), Instagram (Instagram, 2018), and YouTube (YouTube, 2018). An increasing number of people start to engage with live streams (Restream Team, 2021). Facebook reported that more than 800 million people engage with live-streaming content, including workout classes and concerts, every day (Facebook, 2020). The concurrent viewer count of a single channel can range from a few people to millions (Raman et al., 2018; Sheng & Kairam, 2020). Though most live streams have fewer than 100 viewers (Seering et al., 2018), large-scale streams still became a common thing. For example, on April 18, 2020, Nielsen estimates that around 21 million people watched the One World: Together at Home concert across 26 networks, and there were at least 4.9 million interactions across Facebook, Instagram, and Twitter about this event on that day (Nielsen, 2020).

As the audience size increases, the vast number of messages received at nearly the same time becomes a source of “breakdowns” which transform meaningful discussion into a “roar of the crowd” where the chats scroll up too fast to read (Hamilton et al., 2014). People reported not only finding it difficult to read the chat messages, but especially hard to interact with others (M. K. Miller et al., 2017). Although some users feel that watching a massive live stream is a compelling experience, most streamers and viewers say that they appreciate a meaningful and constructive conversation (Hamilton et al., 2014; Lu, Xia, et al., 2018).

Communication breakdown in large-scale communication has been reported in previous literature. Jones et al. (Jones et al., 2004, 2008) have studied the phenomenon of information overload in the context of a Usenet newsgroup and an Internet Relay Chat network. They found that as the overloading mass interaction increases, users are more likely to end active participation, and more likely to generate and respond to simpler messages. They also found active participants of a shared public discourse in a chat channel appear to be limited to a certain number, even the number of users in the chat channel increases to hundreds (Jones et al., 2004, 2008).

Researchers and practitioners have explored many solutions to deal with challenges in mass interaction, such as collaborative tagging (Mamykina et al., 2011; A. X. Zhang & Cranshaw, 2018) and proximity-based content filtering (Viegas & Donath, 1999). Technical advances in machine learning and artificial intelligence (AI) open a new door to handle the challenges using algorithmic approaches. AI has been used in CMC in the domain of machine translation (B. Xu et al., 2014), instant messaging (Hohenstein & Jung, 2018), content filtering (Newton, 2016), etc. An intelligent system equipped with machine learning algorithms or artificial intelligence (AI) can automatically recommend relevant interesting people or content based on users' preferences (Guy, 2018). AI can not only mediate the communication content (J. Hancock et al., 2020), but also operate communication dynamics by selecting interlocutors of a conversation.

Based on the practices and challenges in CMC literature above, I pose three research questions that focus on three sets of activities: understand, design, and evaluation:

1. What motivates people to watch live streaming on SNS? How do viewers engage with the streamer and other viewers in live streams on SNS? What challenges do viewers face when trying to engage with the streamer and other viewers and what causes them?
2. Taking viewers' experiences and opinions into consideration, how can we design interactive applications to facilitate communications in large-scale live streams? How do viewers think of such a system?

3. How can AI be embedded in a live stream system that helps viewers communicate in real time? How do viewers think of such practice? What benefits and risks does AI bring?

Research Goals and Approach

This dissertation aims at exploring design opportunities to facilitate large-scale communication. More specifically, I focus on applying algorithmic approaches in real-time text chat systems and designing intelligent interfaces in the context of live streaming. Live streaming is selected as the research context for the following reasons. First, live stream is a promising technology for supporting real-time communication at distance. It is unique among CMC tools as it integrates multiple communication mediums together and offers rich interactions. Second, the challenges brought by the communication scale and rich interactions, which will be demonstrated in later chapters, are also a challenge for human capacity. Therefore, live streaming may be a suitable experiment field for AI to get involved in human communication. Third, the increasing use of live streaming in recent years indicates the potential that this work can be applied to real-world scenarios.

I apply behavioral social science research methods such as interviews and lab experiments. In addition, I follow a human-centered design approach (Boy, 2011; Cooley, 2000) where the design is informed by user research and evaluated with human participants. I first conduct a qualitative interview study to understand the motivations, engagement, and challenges of live stream viewers. Findings show that though viewers feel a sense of community when watching live streams, they still face challenges such as distraction, lack of direct communication, lack of a friendly environment, and lack of meaningful conversations. Based on these findings, I brainstorm and propose potential solutions that generate subgroups based on similarity. I evaluate the proof of concept in a controlled Wizard-of-Oz lab experiment. The results indicate subgrouping is an effective approach to make chat easier to read and conversation more enjoyable; participants report more interpersonal attraction in similarity-based grouping than random grouping. Then I design and develop Chatbuddies, a new live stream chat framework that creates subgroups and visualizes grouping results plus grouping strategies. The evaluation with real Twitch users in an online experiment shows positive results towards the proposed system while further design consideration and discussion is needed.

Research Contributions

This dissertation contributes to the fields of human-computer interaction (HCI), computer-mediated communication (CMC), and AI-embedded communication. The contributions are summarized as follows:

1. A qualitative study of how people engage with live streams on SNS that identifies two engagement styles: loyal followers and community players, and corresponding challenges for each viewer type.
2. An in-depth review of similarity-attraction effects, including literature and critics, a discussion of underlying factors that influence the similarity-attraction link, and a review of its potential application in CMC.
3. The creation of a novel CMC tool with similarity-based subgroup generation algorithms and algorithm explanation interfaces for large-scale live streams.
4. Evidence for the potential value of the abovementioned tools through a Wizard-of-Oz controlled lab experiment and an online experiment.
5. Design implications for future tools designed for large-scale communication and AI-embedded CMC.

Dissertation Overview

The remainder of this dissertation is structured as follows. In Chapter 2, I review related literature on large-scale communication, the challenges, and potential solutions. Chapter 3 describes an exploratory study on people's motivations for watching live streams, their engagement with others, and the challenges they face. Chapter 4 presents a proposed method that creates subgroups in live streams and a Wizard-of-Oz lab experiment. Chapter 5 describes the design and evaluation of Chatbuddies, a novel live stream chat tool that generates subgroups using clustering algorithms. Finally, Chapter 6 summarizes the results of above studies and discusses the design considerations for similarity-based subgrouping in live streams and design implications for AI-embedded CMC systems.

CHAPTER 2

BACKGROUND

In this chapter, I review previous work, summarize my research motivations, and present the theoretical foundations of potential solutions. I first describe the characteristics of live streaming in real-time text communication and online communities. Next, I summarize the challenges of conversation in large-scale live streams. Then I discuss existing solutions and investigate a potential solution, similarity-affect theory, by reviewing its theoretical foundations and empirical studies in both social psychology and communication areas. Further, I envision the future of the proposed solution to be implemented in an intelligent system using algorithmic approaches and discuss relevant applications. In the end, I summarize the literature and pose my research questions.

A Brief History of Large-scale Communication

Early practices of large-scale interaction can be traced back to Usenet and Internet Relay Chat (IRC). Usenet, one of the earliest newsgroup conversational applications, experienced a course of development since it was created in 1988 by two graduate students in Duke University (Pfaffenberger, 2003). It allows users to read and post messages in topically organized forums, so-called newsgroups. In the beginning, the transmission of posts took hours, and the communication was asynchronous. But soon Usenet grew from 2 articles a day posted at three sites in 1979 to 1800 articles a day posted at 11,000 sites by 1988 (Hauben & Hauben, 1997). In 1996, there were over 17,000 newsgroups and approximately 3 million users (Whittaker et al., 2003). Usenet retired in 2010 because of low usage and rising cost. But the concept of mass interaction remains in other forms of communication.

Almost at the same time when Usenet was created, Internet Relay Chat (IRC) was created by Jarkko Oikarinen in Finland around 1988 (Latzko-Toth, 2010). It is a text-based client-server chat system where users can connect to a server via the Internet from their local client program. Messages are sent and received in real time. IRC was first tested on a single machine with fewer than twenty users. After it was connected to national Internet, it quickly spread to more than 20 countries in America, Europe, Asia, and Australia (Reid, 1991). The scale of IRC services also increased dramatically. According to some reports retrieved from early websites (Latzko-Toth, 2010), around the middle of 1989, there were only 40 servers and an average of total of 10 concurrent users at peak hours. It grew to 211 servers and 20,000 regular users in 1993. In 2011, it peaked at 65,000 users in 40,000 channels (Stenberg, 2021).

Beyond text, audio- and video-based communication became prevalent in the past decade. Discord is a communication platform originally designed for gamers. It offers voice channels that allow a great number of users to have seamless communication while playing games (Ravenscraft, 2020). Clubhouse, an audio-based social media where people can spontaneously jump into voice chat rooms together, exploded in popularity during the COVID-19 pandemic. Leading by celebrities, some chatrooms quickly hit the cap of 5,000 concurrent listeners and overflowed to other social media platforms (Newton, 2021). During the COVID -19 pandemic, working from home became a new normal (Bajarin, 2021). In-person meetings, classes, conferences, and events were replaced by video conferences (Foramitti et al., 2021; Helsen et al., 2021).

Overview of Live streaming Service on Multiple Platforms

Live streaming has become increasingly popular in the past decade (Tang et al., 2016). The content and categories of live streams vary across platforms and user groups. Video gaming is one of the most popular categories of live streams (Hamilton et al., 2014; Kaytoue et al., 2012). People also broadcast and watch creative live streams (Faas et al., 2018; Fraser et al., 2019), talk shows or performances (J. Li et al., 2019), knowledge sharing streams (Y. Chen et al., 2020; Lu, Annett, Fan, et al., 2019), online shopping (Cai et al., 2018), outdoor activities (Lu, Annett, & Wigdor, 2019), etc. Common live stream platforms in North America include early mobile applications such as Meerkat and Periscope (Tang et al., 2016), general live stream platforms such as Twitch (Hamilton et al., 2014), social media platforms such as Facebook and Twitter (Haimson & Tang, 2017). People often share daily life activities and events with friends on Facebook and Instagram, and influencers interact with subscribers on Instagram, Twitter, and YouTube (Raman et al., 2018). Some businesses and organizations share livestream events with the public on their Facebook page or YouTube channel (Luo et al., 2020).

By integrating live video, text chat, and a variety of other tools, live streams offer rich ways for people to interact with each other. Viewers can chat with the streamer, ask questions, request activities, send digital gifts, or donate money (J. Li et al., 2019; Lu, Xia, et al., 2018; D. Wang et al., 2019). They can also chat with other viewers, especially when the streamer is not able to respond in time (Lu, Annett, Fan, et al., 2019) or when the stream is for relaxation and social purposes (Taber et al., 2020). A streamer may read the comments from viewers, answer their questions, and even invite guests to co-broadcast the stream (J. Li et al., 2019).

Real-time Text Communication in Live Streams

Text chat is an essential component of live streaming (Hamilton et al., 2014; Lu, Xia, et al., 2018; Tang et al., 2016), allowing viewers to communicate with the streamer or other viewers. Typically, it is displayed as a chat window to the side of a live video. Most platforms support chat features such as emoticons. Some support images, stickers, and donations (Lu, Xia, et al., 2018; Tang et al., 2016; D. Wang et al., 2019). Although the features vary across different platforms, the overall display and functionality of live chat are similar to Internet Relay Chat (IRC), where people can communicate via text messages in a “chat room” or a “channel” (Hamilton et al., 2014). The size of the channel varies from a few to a million, depending on the number of concurrent viewers (Sheng & Kairam, 2020).

Early work in CMC investigated the characteristics of text-based multi-party communication. Herring (1999) pointed out two problems of text chat: lack of simultaneous feedback and disrupted turn adjacency. The absence of visual and audio cues makes it challenging for people to take turns in an orderly fashion. Multiple users may send messages at the same time or hold overlapping conversations on different topics. Users have developed strategies to overcome these barriers, such as the use of turn change signals, cross-turn reference, and organization of topics by channels or threads (Herring, 1999). However, many of these techniques don’t scale well to very large audiences. As more people join the IRC channel, chatters see multiple conversation threads including those they are not engaging with. Some chatters engage with multiple threads at the same time, it becomes difficult to identify which thread they are responding to.

How Live Stream Chat Differs from Other Text-based CMC

Live stream chat differs from other text-based CMC in several ways. The first and most salient difference is that live stream chat belongs to a class of technologies such as social TV (Harboe et al., 2008) or live-Tweeting (Schirra et al., 2014) that allow people to do things and chat at the same time. People comment on the ongoing event, share thoughts, and discuss with each other in chat rooms or on Twitter using hashtags. Second, the maximum scale (the number of people in the chatroom at the same time) of live stream chat is much larger than most text-based chat tools. Jones et al. (2008) conducted a study on the AustNet IRC network and found the maximum number of users active in one chat room was under 300, whereas in many livestream events, the number of concurrent viewers can easily exceed 1 million (Nielsen, 2020; Raman et al., 2018). Third, the purpose of sending chat and the chat content is somewhat different in live streams and text-based CMC platforms. For example, live stream chats are often dominated by the events in the live video (Recktenwald, 2017); repetition and "shout-outs" are common in live stream chats (Ford et al., 2017; Nematzadeh et al., 2019), but these phenomena were not reported in previous studies on IRC (Herring, 1999; Werry, 1996).

How Live Chat Affects Viewership

Research has shown that live chat affects the emotion, behaviors, and opinions of concurrent viewers (J. Guo & Fussell, 2020; Luo et al., 2020; Maruyama et al., 2014, 2017; Seering et al., 2017). Evidence of emotional amplification and emotional contagion has been reported in previous studies (J. Guo & Fussell, 2020; Luo et al., 2020). Researchers also reported that people's behaviors in live stream chats are consistent with theories of imitation, confirmation to authority, and deterrence (Seering et al., 2017). They investigated three types of behaviors, smile, spam, and question. They found that if a message included any of the three behaviors, the following ten messages had significantly more behaviors of the same type than if the message did not have a behavior. The impact of chatters with higher social status such as moderators is stronger than regular users. Deterrence theory argues that individuals' behaviors will change under the threat of punishment. The authors found if a certain type of behavior (e.g., smile) was banned, there were fewer such type of behavior in the following ten chats than if this behavior was not banned. This research found evidence of social science theory on live stream chat through a quasi-experiment, but it did not measure viewers' subjective attitudes and opinions towards the impact of others' chats.

In a survey study where participants reported their affect, knowledge, memory, and vote choice before and after watching a political debate and participating in live tweeting, researchers found that active tweeters changed their vote choice to reflect the majority sentiment on Twitter, compared with tweet observers and no tweet participants (Maruyama et al., 2014). A similar conformity to the majority phenomenon was reported in another study where participants watched a political live stream with live comments on the side. In addition, participants reported thinking more about the discussion before and after they posted if they saw positive opinions (Maruyama et al., 2017). This study focused on a special type of live stream, political social TV, where the discussions were generally elaborative, the conversation flow is different from other types of live streams such as gaming and talk show.

Challenges of Real-time Communication in Large-scale Live Streams

There are two major challenges in large-scale live stream text chats: making sense of the conversations and participating in meaningful interactions.

Text-based group chat has a long history of information overload problems. Jones et al. (2008) used an information-processing constraints (IPC) model to explain the dynamics of online communication. They found chatters tend to become less active and send simpler messages as the number of messages increases (Jones et al., 2004). In gaming live streams, the phenomenon of “copypasta” has been frequently reported. Even in small-scale live streams, viewers found it hard to catch up with the events in the streams or the chats.

Previous studies reported that larger group does not support interactions (Kim, 2013) Larger groups attract more members than smaller groups but also lose more. Hamilton et al. (2014) noted that information overload in live stream chat can break down meaningful conversation and constrain further interaction. Nematzadeh et al. (2019) analyzed a large dataset from Twitch and found a transition from a conversational state to a state with more repetition and less information per message. The definition of "meaningfulness" varies in different communities. For example, in gaming communities, the seemingly chaotic and repetitive chats may create a sense of community (Hamilton et al., 2014). Some researchers claim that the "practices of coherence" make massive chats legible and meaningful (Ford et al., 2017). However, in other contexts where people have more serious motivations such as learning new things, exchanging opinions, and making friends, the problem caused by overwhelming numbers of messages still exists and can reduce people's enjoyment of their live streaming interactions.

Online Communities Formed in Live Streams

Like other online spaces, many live streams serve as a virtual place for the streamer and viewers to emerge, socialize, and participate. Hamilton et al. (2014) conducted an ethnographic study on Twitch.tv, one of the most popular live stream platforms at that time, and reported that people engaged in Twitch live streams for two reasons: the unique content of a particular stream and their interests in that stream's community. Live stream viewers develop shared history over time as they frequently visit the site and interact with each other. Ephemeral in-game events and emotional reactions expressed by a huge volume of chats contribute to constructing a sense of community and the feeling of being together.

Besides gaming communities, other categories of live stream communities have been studied, such as creative activities (Faas et al., 2018; Fraser et al., 2019), programming (Y. Chen et al., 2020), cultural heritage (Lu, Annett, Fan, et al., 2019), and outdoor events (Lu, Annett, & Wigdor, 2019).

Creative live streams generally have a smaller audience size compared with gaming streams. Fraser et al. (2019) reported that viewers watch creative streams for learning, community, and entertainment. Unlike lectures and tutorials, live streams offer direct interaction with the streamer and other viewers, which cultivates a "mentorship community" of people with similar interests (Faas et al., 2018; Fraser et al., 2019). Some communities extend the interaction beyond the in-stream live chat to off-stream asynchronous communication such as Discord, an instant messaging and digital distribution platform. The extended interaction helps community members to stay in contact and strengthen relationships (Faas et al., 2018).

Live stream communities in other demographics have also been studied by HCI researchers. Lu et al. (2018) explored live streams in China through a mixed-method study. They highlighted reward-based systems and fan groups on instant messaging applications as different practices compared with live streams in North America. Streamers create fan groups on Wechat or QQ, which are popular instant messaging apps that support multiple modalities such as text, emoji, stickers, voice, image, and videos. Streamers use fan groups to send notifications about going live, ask for feedback about certain streams, or ask the viewers about their interests. Viewers join fan groups to learn more about the streamer and make friends. Some viewers disclose personal information and even meet up offline. Fan groups serve as a close-knit subgroup inside a large live stream community around the streamer. Live streaming services have been used in rural areas to accommodate the needs of elderly villagers and support the co-presence of local church communities (Struzek et al., 2020).

Sheng & Kairam (2020) investigated how relationships formed on Twitch and how the relationships grow stronger over time. They reported that the affordances of Twitch, such as the absence of audiovisual cues, facilitate self-disclosure and emotion expression (Walther, 1996). The low initial barriers of connecting to strangers encouraged people to express their "true self" and make it easier to share embarrassing, controversial, or upsetting material. The lack of audiovisual cues also shapes how people receive others' self-disclosure such as less judgement on others' physical appearance. Authors also found that assumed common ground fostered by shared context and interests, for instance, interests in a specific video game that they do not largely share in the general population, allows their conversations to delve deeper when they shift from general Twitch chat to more intimate modes of communication. Consistent with previous research on online communities (Kraut & Resnick, 2012), the relationship grows stronger when the members have frequent interactions across multiple modalities such as text chat, voice call, and video chat.

Online Communities Formed in Social Watching

Social watching, also called social TV and second screen, is a special phenomenon in live broadcasting, where viewers actively communicate via social media such as Facebook and Twitter during watching television (Maruyama et al., 2017). Though social watching does not fall into the definition of live streaming studied in this dissertation, it is worth reviewing how people find communities and get connected while watching videos together. People turn to the second screen to discuss the broadcast content, learn further information, be part of a community, and connect with people who share the same interests (Gil de Zúñiga et al., 2015; Schirra et al., 2014).

Maruyama et al. (2017) conducted a controlled lab experiment where participants watched a simulated video and saw others' comments. They designed a 3 x 3 factorial design the factors are interactivity (post comments and receive feedback, post comments without feedback, no interaction) and post opinions (support, opposite, balanced). They found that people who received positive feedback experienced stronger group membership, need fulfillment, mutual influence, and emotional connection. Further, the more people felt they belonged to the community, the more they elaborated before sending messages in the conversation.

How Live Stream Communities Differ from Other Online Communities

Kraut et al. (2012) summarized three kinds of commitment in online communities: 1) affective commitment, or how much an individual wants to continue as a community member based on closeness and attachment to the group or group members; 2) normative commitment, or how much an individual believes they are ought to stay or are obligated to stay; 3) need-based commitment, or how much an individual believe they will obtain more benefits if they stay in the community than be out of it. Members in most online communities experience all three commitments while the extent varies. Based on previous literature and my own observations, most communities formed in live stream contexts rely on members' affective commitment. Members in special groups such as small streamer mutual help groups may have more need-based commitment. Moderators may have more normative commitment as they already built personal connection with the streamer or were requested to be a moderator (Wohn, 2019).

Social psychologists distinguished two bases of affective commitments (Kraut & Resnick, 2012). One is identity-based commitment, a feeling of being part of the commitment and helping to fulfill its mission, in another word, attachment to the community as a whole. Another is bond-based commitment, feeling close to specific members of the group. These attachments have been reported in previous literature that people engage in live streams because they like the atmosphere of the community, or they have developed close connections (Hamilton et al., 2014; Sheng & Kairam, 2020).

In addition to affective commitment, communities formed in live stream context have two other salient characteristics: synchronicity and multi-modality. Communication in many online communities is asynchronous, such as Reddit and Wikipedia Editors. Live streaming itself is a synchronous CMC service, members know who is also watching the stream, and they can get chat responses in real time. Synchronicity is not limited to real-time communication but the status of "doing things together" (Schirra et al., 2014). Like in-person activities such as concerts and sports events, "being together" brings participants a feeling of connectedness and excitement (J. Li et al., 2019; Musabirov et al., 2018). Besides broadcasting on Twitch, many streamers use other CMC applications to manage the community, such as Discord servers, WeChat, and QQ groups, so that community members to hang out when the streamer is offline (Faas et al., 2018; Lu, Xia, et al., 2018). Though live streaming is a major virtual space for people to hang out, people tend to prefer the flexibility of other CMC applications that offer multiple modalities such as voice, stickers, and tips (Lu, Xia, et al., 2018).

Challenges of Friendship Development in Large-scale Live Stream Communities

Literature shows that some people engage with live streams to make friends, extend social connections, and find a community (Hamilton et al., 2014; Hilvert-Bruce et al., 2018; Lu, Xia, et al., 2018; Sheng & Kairam, 2020). As I discuss above, most members of communities formed in live streaming context have strong affective commitment, either identify- or bond-based. Identify-based commitment refers to the attachment members have towards the group as a whole; whereas bond-based commitment describes the attachment members have towards particular members in the community. Making friends with individual community members is a process of constructing bond-based commitment. There are two major challenges in friendship development in large-scale streams: finding potential "friends" and strengthening connections.

Members in a large live stream community should share at least one common ground: interest in the content of this live stream channel. Despite this common interest, apparently people cannot make friends with all other community members. Lots of live stream viewers are willing to befriend and communicate with strangers. They develop interpersonal relationships with those who share some other similarities, such as geolocation, mutual friends in real life, and other hobbies or interests. Or they become friends with those they see or interact the most (Lu, Xia, et al., 2018; Sheng & Kairam, 2020). People have limited capacity to process information in a large-scale live stream chat. As a result, they are not likely to read through all the messages or recognize all the chatters. It takes time and effort to recognize someone they may be interested in and want to interact with. After initial interactions, dyads need repeated exposure and social interactions to strengthen the relationship. In a large community with hundreds of members, not all members have equal visibility (Sheng & Kairam, 2020). Previous literature reported a stronger association between social motivators and live stream engagement for viewers who mostly watch small channels than large channels (Hilvert-Bruce et al., 2018). Small channels provide richer and more direct interactions, more frequent exposure, and higher visibility of individual members. As the community size grows, although the potential communication partner increases, it is hard to have one-one conversations (Hamilton et al., 2014). The increase of chat volume even has negative effects on perceived interactivity and enjoyment were reported in past research (Haimson & Tang, 2017).

Potential Solutions of Challenges in Live Stream Chat

Researchers and practitioners have implemented a variety of solutions to handle chat message overload, such as limiting the number of messages sent by each audience member (slow mode[1]), restricting the chatters to a certain number of people (follower-only[2] and subscriber-only), or filtering messages to display the most relevant ones (Koroleva & Bolufé Röhler, 2012). A

[1] Slow mode: streamer or moderator can set a limit on how often users in the chat room are allowed to send messages.
[2] Follower-only: only users who have followed for a specific amount of time can chat

limitation on messages or people is a straightforward solution and some strategies have positive effects on moderation (Petrocelli, 2017) , but it constrains some viewers' role to be observers instead of participants. Filtering is commonly used in SNS, especially for moderation and personalized recommendation. But many existing filtering models tend to consider the filtered content as a unidirectional, asynchronous information source instead of part of a bidirectional interaction.

Van Alstyone and Brynjlfsson (1997) argued that virtual villages and online communities would eventually balkanize towards specialized groupings to cope with the plethora of groups and millions of internet users. Viegas and Donath (1999) presented Chat Circles, an abstract graphical interface that used a physical proximity metaphor to break large groups into clusters. Miller et al. (2017) proposed a similar idea in live streaming: that the system can assign audience members to random positions and display messages based on a user's "proximity" and the number of upvotes (a "neighborhood"). Their evaluation shows that the system helps users read and understand messages and it highlights important content in the chat channel. Although the "neighborhood" is formed around individual users, which means every user has a unique neighborhood, the authors claimed that in practice it will be symmetric for most pairs of users.

Randomization is a convenient and efficient approach to solve the problem of information overload. It has some trade-offs too. Random grouping overlooks viewers' interaction history and conversation interests. It solves the problem from a utility point of view, but it can't help develop or maintain underlying social connections (Hamilton et al., 2014). Research shows that bond-based attachment (attractiveness of individual group members) is crucial to the maintenance of online communities (Farzan et al., 2011; Kraut & Resnick, 2012).

Existing solutions, their strengths and weakness, and my previous experience as a designer triggered me to think, is there any way that I can not only solve the problem but also explore a new pattern of communication? How might I think of this issue not solely from the perspective of problem-solving, but from the angle of experimental innovation? What social psychology and communication theories can be applied to the design of a live streaming service?

Kraut et al. (2012) summarized four factors that community designers can harness to affect members' interpersonal attraction: repeated exposure, similarity, social interaction, and self-disclosure. Similarity is a keyword that repeatedly appears in social psychology literature. The phenomenon that people tend to be attracted by those who share similar personalities or attitudes, or called "similarity-attraction effect", has been studied for decades (Berscheid & Reis, 1998). What will happen if I apply similarity-attraction theory on live stream chat? Can it help with the challenges I discussed above?

In the next section, I elaborate on the development of similarity-attraction theory in both social psychology and communication areas, critiques it encountered, and its applications in CMC.

Similarity-Attraction Effect

The earliest work of similarity-attraction effect can be traced to Newcomb (1961)'s longitudinal study of friendship formation in college dormitories and Byrne and his colleagues' series of laboratory experiments (Byrne, 1961, 1971; Byrne et al., 1970; Byrne & Nelson, 1965). It refers to "the widespread tendency of people to be attracted to others who are similar to themselves in important respects. Attraction means not strictly physical attraction but, rather, liking for or wanting to be around the person." (Baumeister & Vohs, 2007).

The Bogus Stranger Paradigm

Byrne and his colleagues conducted a series of studies and proposed that attraction is a linear function of attitudinal similarity. They developed an experiment procedure, "bogus stranger", to investigate the effects of similarity on interpersonal judgement. In the study, participants first complete a 26-item self-description survey about their attitudes on a variety of dimensions (e.g., God, premarital sex relations, western movies). Next, they were shown responses from a bogus stranger. Then they were asked to evaluate the other in terms of intelligence, knowledge of current events, morality, and adjustment (Interpersonal Judgement Scale, Byrne, 1971) plus two attraction items, how well they felt they would like this person and whether they believed they would enjoy working with him (or her) as a partner in an experiment. Results indicate that a similar bogus stranger received significantly more positive feelings than a dissimilar other.

Later Byrne and his colleagues explored similarity effects by varying similarity dimensions such as spending money (Byrne et al., 1966) and personality characteristics (Byrne et al., 1967), stimulus modes (Byrne & Clore, 1966), population (Byrne et al., 1969; Byrne & Griffitt, 1966), and attraction measurements (Byrne et al., 1971). Results of these studies consistently support the assumption of positive association between similarity and attraction.

Actual Similarity, Perceived Similarity, and the Role of Communication

Despite the overwhelming supportive evidence of the similarity effect, critiques have been raised regarding the integrality of the effect. Byrne (1971)'s "bogus stranger" technique has been criticized for a long time because of its low ecological validity (Huston & Levinger, 1978; Sunnafrank, 1992). It was uncommon in daily life that an individual can receive a bogus stranger's response to a self-descriptive survey in the 1960s when Byrne's study was first published. It may be different today since many relationships start on the Internet through SNS, dating apps, and chatrooms (Finkel et al., 2012; Rosenfeld & Thomas, 2012). Summarized information is often accessible online prior to or immediately after the first interaction (Finkel et al., 2012; Sprecher, 2014), such as a profile page presenting one's location, education, experience, and hobbies (Lampe et al., 2007).

But communication is essential for relationship initiation and development (Sprecher et al., 2008). Evaluating others based on questionnaire responses without any interaction does not have much practical significance if the goal is to understand communication and social interaction. Based on a few studies that involved interaction (Brewer & Brewer, 1968; Byrne et al., 1970; Levinger, 1972), Sunnafrank and his colleagues (Sunnafrank, 1983, 1984, 1985, 1986; Sunnafrank & Miller, 1981) conducted a systematic line of similarity-attraction research that introduced a get-acquainted conversation in the experiment procedure and proposed a different argument that attitude similarity and attraction are unassociated in beginning communicative relationships. In one early study by Sunnafrank and Miller (1981), instead of a bogus stranger, participants were paired with another same-sex stranger. In the first step, participants filled out a two-item dichotomous measure about their attitude toward nuclear power plants and preparedness for war. Then they were given their partner's response and were asked to form some opinions of the other. Participants were randomly assigned to interaction and no-interaction conditions. In interaction condition, participants had an initial conversation for 5-minutes, whereas in no-interaction condition, participants did not meet each other. In the end, they were given an interpersonal attraction scale adapted from Byrne' Interpersonal Judgement Scale (1971). The results of no-interaction condition were consistent with Byrne's study (1961), however, no difference in attraction among similar and dissimilar interactants was found. Both similar and dissimilar conversational partners were equally attracted to the other.

Later other factors were investigated, such as the introduction of both initial conversation and attitudinal topic conversations (Sunnafrank, 1983, 1984), opposite-gender pairs (Sunnafrank, 1985), and perceived similarity (Sunnafrank, 1986). Sunnafrank reported that the effects of manipulated similarity (the two attitudinal questions used in the study) only happened in the pre-conversation stage; actual attitude similarity had no association with attraction after a getting-acquainted conversation. Furthermore, an initial conversation even had a positive influence on attraction for dissimilar strangers (Sunnafrank, 1983). There is one exception where the effect of actual attitude similarity remains intact such that participants evaluate the other after an attitudinal discussion instead of a get-acquainted conversation (Sunnafrank, 1984).

Sunnfrank's measurements of attitudinal similarity may be too simple compared with the 13-item scale used by Byrne (1961). Sprecher (2014) modified Sunnfrank and Miller (1981)'s method on similarity measurement, conversation content, and dependent variables. In her study, similarity is measured by the proportion of agreed items on a 17-item dichotomous scale with innocuous questions including personality, hobbies, attitudes, etc. A free get-acquainted conversation was replaced with a structured self-disclosure conversation including values and past experiences (Aron et al., 1997). New factors such as enjoyment of interaction and closeness were added as dependent variables. Results showed that perceived similarity was a strong indicator of liking and other interpersonal judgement. Actual similarity, however, did not affect liking and other outcomes once the participants interacted. The same result of actual and perceived similarity were reported in another study where participants discussed political views, leisure interests, or just got acquainted (Sprecher et al., 2015). In a study investigating similarity-attraction link in existing relationships, the effect of perceived similarity was reported to be mediated by consensual validation, certainty of being liked, enjoyment of the interaction, and self-expansion (Sprecher et al., 2013).

Besides short interactions between strangers in a lab setting, studies on existing relationships also reveal that perceived similarity has a significant effect on interpersonal attraction (Strauss, 1993). A meta-analysis examined 313 studies and reported manipulated actual similarity is important in no-interaction and short-interaction experiments but not significant in existing relationships. Perceived similarity, in the contrast, predicted attraction in no-interaction, short-interaction, and existing relationships (Montoya et al., 2008).

However, supportive evidence of the association between actual similarity and attraction involving communication also exists. These studies have the following study design that is different from those described above: a) participants discussed topics that closely relevant to the similarity measurement (Brewer & Brewer, 1968; Sunnafrank, 1984); b) similarity was measured with a more granular scale, e.g., 40-50 items (Byrne et al., 1970; Griffitt & Veitch, 1974); c) participants were not informed of the specific items in similarity measurements but a comprehensive summary of their partner. For example, a proportion of similar responses in Byrne, Ervin, and Lamberth (1970); a general description such as “similar” or “different” on a variety of dimensions in Cappella & Palmer (1990). Some studies did not inform other participants’ attitudes at all (Griffitt & Veitch, 1974; Neimeyer & Mitchell, 1988).

Because of methodological ambiguity, contracted results of similarity effects in the context of communication come as no surprise. The above-mentioned studies were different in terms of similarity measurements (e.g., multiple-item or dichotomous attitude scale), information delivery approaches (e.g., read the other’s survey responses or inform the proportion of identical responses), communication content (e.g., 5-minute get-acquainted chat or 20-minutes discussion on a certain topic), and communication instructions (e.g., structured or free-style, taped or non-taped). Environmental factors such as room temperature and background music may have effects on the result too (Byrne, 1992; Montoya et al., 2008). I summarized study design, measurements, and results of some lab experiments in Table 1.

In this dissertation, similarity-attraction effect is studied as a potential solution of challenges in large-scale real-time communication. A justified and theoretically sound design requires that factors that may influence the effects of similarity on interpersonal attraction be inspected and discussed. In the next section, I want to have an in-depth discussion on the involving factors from the aspects of similarity dimension, similarity information delivery, information exchanged in the interaction, and type of relationships.

Similarity Dimensions

People can be similar to one another on a number of dimensions, including attitudes and beliefs (Byrne, 1961), personality traits (Banikiotes & Neimeyer, 1981; Byrne et al., 1967; Duck & Craig, 1978), hobbies (Curry & Emerson, 1970; Werner & Parmelee, 1979), and background or demographic characteristics (Baxter & West, 2003; McCroskey et al., 2006; Sprecher, 1998).

There are many ways to define the similarity between two individuals. Most similarity-attraction studies employed attitudes to represent similarity. According to Clore & Byrne (1974)'s reinforcement framework, people with similar attitudes are seen as rewarding because they can validate one's ideas and attitudes, and thus reinforce the logic and consistency of the worldview. Later work expanded the attributes to other dimensions such as personality traits, interests, and interaction styles. Here comes a question, does a similarity dimension matter? Does one dimension (e.g., attitudes) have a more significant effect than other dimensions?

Publication	Actual Similarity Measurement	Perceived Similarity Measurement	Similarity Manipulation	Information Exchanged	Actual similarity Affects Attraction?	Perceived Similarity Affects Attraction?
Brewer and Brewer, 1968	Attitudes (20 items)		General description (high or low compatibility)	20-min discussion on attitude topics (transcribed)	**Yes**	N/A
Byrne, Ervin, and Lamberth, 1970	Attitudes (50 items)		Proportion of similar items	30-min F2F "coke date"	**Yes**	N/A
Sunnafrank and Miller, 1981	Attitudes (2 items)		Survey response	5-min get-acquainted chat (audiotaped)	**No**	N/A
Sunnafrank, 1983	Attitudes (2 items)		Survey response	5-min get-acquainted chat or get-acquainted plus attitude topics, audiotaped)	**No**	N/A
Sunnafrank, 1984	Attitudes (2 items)		Survey response	5-min get-acquainted chat or attitude topics or both (audiotaped)	**Partially**	N/A
Sunnafrank, 1985	Attitudes (2 items)		Survey response	5-min get-acquainted chat (avoid opinion topics, audiotaped)	**No**	N/A
Sunnafrank, 1986	Attitudes (2 items)	Predicted partners' response on 13 attitude topics from Byrne (1971)	Survey response	5-min get-acquainted chat (avoid opinion topics, audiotaped)	**No**	**Yes**
Cappella and Palmer, 1990	Attitudes (50 items)		General description (similar or dissimilar)	30-min free chat (videotaped)	**Yes**	N/A
Tidwell et al., 2013	personality, political conservation, traditionalism, etc.	General (2 items) and trait-specific questions (1 item)	N/A	4-min speed dating	**No**	**Yes**
Sprecher, 2014	Attitudes, personality, hobbies, interests, etc. (17 items)	General question (2 items)	Survey response	5-min structured self-disclosure task with two rounds	**No**	**Yes**
Sprecher et al., 2015	political views (4 items), leisure interests (19 items)	General (2 items) and topic-specific questions (1 items)	N/A	15-min F2F discussion (politics, hobbies, get-acquainted)	**No**	**Yes**

Table 1. Literature of Similarity-attraction Effect with Communication

Neimeyer and Mitchell (1988) conducted an eight-week study to examine the discriminability of attitude, personality, value, construct, and structural similarity as prediction of attraction. They found attitude similarity was a significant predictor of initial attraction across stranger pairs in the first week, whereas only personality and structural similarity predicted later attraction. Even subtypes of personality affect attraction at different stage of a relationship. Duck and Craig (1978) reported that easily accessible personality traits have more influence on early acquaintance, whereas similarity on fundamental traits have more influence on established relationships.

Treger and Mascialebc (2018) surveyed 250 Amazon Mechanical Turk (MTurk) workers about their preference for 18 similarity domains on three types of relationships. The most important domains were reported as political views, career goals, food preferences, travel desires, and music preferences. The fact that some dimensions received more preference than others indicates that ***the importance of similarity on a certain*** *dimension* should affect the attraction in addition to similarity itself. However, most studies neglected this variable but assumed the importance of attitudes.

Little research has been done to investigate the preferences for similarity dimensions on relationship developed online. Evidence in previous social psychology studies that studied offline interaction may not be useful in an online environment, as people have different motivations and approaches when developing relationships online (Chan & Cheng, 2004; Scott et al., 2007). It would be wise to dig literature on online relationship development to find the answer.

Similarity Manipulation

In many studies participants received pre-acquaintance information about the other. The most used method is to present the participant a survey response of the other as in Byrne (1961), Sunnafrank (1983) and Sprecher (2014)'s studies. In other studies participants were given a proportion of similar responses (Byrne et al., 1970) or a general description such as "someone who has opinions, beliefs, attitudes, and values that are quite similar to your own" (Cappella & Palmer, 1990). Yet few studies have specifically compared the effects of similarity information delivery approaches, it seems logical to hypothesize that a more generalized description of similarity would have more manipulation effect on global perceived similarity, compared with a list of survey responses. The survey items may not be seen as important in friendship or other relationship formation, for example, opinions on nuclear power plant in Sunnafrank's studies (1983, 1984, 1985, 1986).

Nowadays people use SNS to self-disclose information such as identity, hobbies and interests (Twitter, Facebook), work experience (LinkedIn), partner expectations (dating apps), etc. Like in the studies where participants read a stranger's survey response, SNS users identify similarities and dissimilarities when viewing other's profile. Some platforms even highlight the similarities such as mutual friends and things in common (Bell, 2017; Nieva, 2018). In large-scale communication, people interact via text chat, audio or video call. The first impression of another user is often the content they speak, their voice or physical appearance, then a hyperlink provides a convenient way to view other's profile. The effects of similarity information delivery methods are under exploration; therefore designers should carefully select an approach that best fits design goals of a new CMC tool.

Information Exchanged in the Interaction

The balance between pre-attained information and information obtained in the process of interaction should have a huge impact on the study results. In previous studies, participants' impression of the other was formed mainly based on the provided information and may be specified with regards to the listed survey responses. In the later interaction, once the participants gained information of the other in other aspects, they would form a new impression and evaluate the degree of similarity from other perspectives. It may explain why the effect of pre-interaction similarity manipulation disappeared when interaction topics were much more open-ended than the pre-acquaintance survey responses (Sprecher, 2014); and why the effect did not disappear in conditions where participants only discussed the topics in the pre-acquaintance survey (Brewer & Brewer, 1968; Sunnafrank, 1984). We can even consider the difference between information gained before and after interaction as a covariate of the similarity effect. In another word, the closer the information gained before and after the interaction, the more salient the similarity-attraction link, though it is difficult to measure.

Conversations in mass interaction are often topicalized, from channels in early IRC (Werry, 1996) to channels in today's live streams (Hamilton et al., 2014). Popular live streams usually have a specific topic or category, such as gaming (Hamilton et al., 2014), creative work (Fraser et al., 2019) and knowledge sharing (Lu, Annett, Fan, et al., 2019), where viewers comments, ask questions, and exchange ideas about the events or topics in the live video. This characteristic provides some insight about what dimension can be used to measure community members' similarity, such as preferences of a game character in a video game community, experience with a domain in a knowledge-sharing community. Members may estimate how similar they are to another by viewing their profiles, which may affect their sequent interaction. Sometimes a live stream may go "off-topic", for example, besides playing video games, a streamer may chat, eat, even sleep in a stream (Lorenz, 2021). As the chats are often about the events happening in the live video (Recktenwald, 2017), the entire community may "go off-topic together" (Kraut & Resnick, 2012) such that members talk about a variety of things, including personal life and self-entertainment (Musabirov et al., 2018; Taber et al., 2020). As people see another side and get richer information about each other, the effects of similarity manipulation may not be salient anymore, as it shows in previous studies (Sprecher, 2014; Sunnafrank, 1983, 1984, 1985, 1986). "Going off-topic together" can build both identity- and bond-based attachment by increasing interpersonal interactions and creating an opportunity that redefines the community's identity separately from its original one (Kraut & Resnick, 2012), which is still the goal of applying similarity-attraction theory.

Type of Relationship

Most lab experiments investigated similarity effects on strangers. Participants were told to get acquainted with another person for the sake of becoming friends (Sprecher, 2014; Sunnafrank, 1983, 1984, 1985, 1986), colleagues (Brewer & Brewer, 1968), or date mates (Byrne et al., 1970; Tidwell et al., 2013). In real life, individuals build a variety of relationships with others, for instance, friends, colleagues, romantic partners, collaborators; in each relationship people have different motivations and expectations. In Treger and Masciale (2018)'s study, participants were asked the degree to which each of the 18 domains is important to them when forming three types of relationships: new friendships, casual/short-term romantic relationships, and long-term romantic relationships. Results showed that participants were most selective for long-term partners, expecting general similarity more than the other two relationships. When evaluating supervisor's attractiveness, students prefer those with higher skills instead of those with higher attitudinal similarity (Hester et al., 1976). While in argumentative discussions, similarity in communication skills is positively associated with participants' ratings of social attraction and competencies on their adversary (Waldron & Applegate, 1998).

Even in the same relationship, people's motivation and practices vary in context and stage. Literature reveals that offline friendships tend to involve more interdependence, intimacy, breadth, and depth than online friendships (Chan & Cheng, 2004; Scott et al., 2007). Similarity effects in friendships also change over time. Literature shows that the effects of attitudes and personal construct similarity on mutual attraction may fade away as friendship develop (McCarthy & Duck, 1976; Neimeyer & Neimeyer, 1983).

Besides factors discussed above, other variables have been reported to mediate or moderate the similarity-attraction effect such as social comparison orientation (Michinov & Michinov, 2011), warmth, competence (Vione, 2016), enjoyment of interaction (Hampton et al., 2018). Hampton et al. (2018) reported that certainty of being liked was the strongest mediator prior to a conversation and enjoyment of interaction was the strongest mediator after a getting -acquainted conversation.

The above discussion covers factors that may influence similarity-attraction effect in both previous studies and potential applications in mass interaction and online communities. Some factors received little examination or have long been neglected. Though the purpose of this dissertation is not to investigate the effects of these factors, it is worth discussing them before coming up with solutions or new design.

Similarity Effect on Group Members

A large body of work inspected similarity-attraction effect in dyad relationship, but few works examined the interpersonal relationship among group members. Modifying Byrne's bogus stranger technique (1961), Good and Good (1974b, 1974a) conducted studies on students' evaluation of association members and occupation employees. They first asked participants to fill out a 14-item attitude scale, then presented them with a summarized response from members of a group (association or occupation) in the format of "Most people engaged in this organization, 15% (85%) believe that money is not one of the most important goals of life." Participants evaluated the entire group from the aspects including how much they like this group and how much they would enjoy entering the group. Consistent with Byrne et al.'s study, results indicated that similarity was associated with the liking of desire to enter. Hansson and Fiedler (1973)'s study on military engineers showed that perceived similarity of personality and values to members in the organization was related to attraction to the organization, but only for relationship-motivated persons not for task-motivated persons.

Applications of Similarity-Attraction in CMC

Similarity-attraction has been wide adopted in many CMC applications, such as friend recommendation and dating apps (Rosenfeld & Thomas, 2012), where users get "matched" with another person. To better understand the effects of similarity on CMC, some researchers and design practitioners proposed CMC systems that connect similar people and evaluated both communication and interpersonal results.

Cosley, Ludford, & Terveen (2003) conducted a study to explore the similarity-effect on task-focused dyad interaction where two strangers played a trivia game and interacted via text chat. Participants were paired based on similarity of interests in task-relevant topics (e.g., popular culture, sports, U.S. history), computed by cosine similarity. Results showed task-relevant interests did not affect task performance, self-rating of collaboration quality, or liking towards the partner. However, demographic similarity affected how people interacted even though it was not explicit in the interaction. Same-gender teams had more social and personal information exchanges. Similar education background pairs talked more, played the game longer, and reported having a better quality of collaboration. It was observed that people actively asked, offered, or inferred personal information during the interaction, even though it was not required. This study focused on dyad and task-focused interactions, the results may extend to group communication and casual interactions.

Ludford, Cosley, Frankowski, & Terveen (2004) conducted a five-week field experiment to investigate the effect of similarity and uniqueness on movie-related topics discussions. The experiment used a 2 x 2 design where 8 groups were formed based on similarity of movie ratings and whether or not receiving weekly emails advertising them of a unique perspective that can be brought to the discussion topic. Results showed that people can identify similar others through anonymous text-based online interaction without prior information; people in similar groups contributed less than those in dissimilar groups. This study introduced a trade-off when designing online communities: contributions versus attraction. Though it did not evaluate the likings toward in-group members, it is reasonable to hypothesize that people favor those who share similar movie tastes versus those who may increase group diversity (Gómez-Zará et al., 2020). It depends on the community designers to decide which design goal they pursue. If the goal is to let people like each other, similarity may be applied to create groups. If the goal is to encourage participation, as the authors suggested, it's better to "favor creating dissimilar groups in situations where disagreement can be tolerated."

Kaptein, Castaneda, Fernandez, & Nass (2014) proposed a when-similar design paradigm in that people discover they are carrying out a specific activity with another individual at the same time. In a study where participants chatted with an ostensible partner through text messages, people in when-similar conditions reported that they felt more social connectedness with the partner. In a follow-up study, participants were shown their ostensible partner's profile including information such as gender, age, academic focus, and three favorite pastimes. Results showed that people tended to like a similar partner more after reading their profile. But the effect of when-similarity was only strong and positive when the partner was initially perceived as dissimilar. In similar demographic condition where participants already established attraction, when-similarity did not increase social connectedness. Time is a unique dimension that has not been widely adopted in CMC applications. With the recent growth of real-time services such as on-demand assistance (Y. Chen et al., 2017), e-counselling (Haner & Pepler, 2016), and live broadcasting (Hamilton et al., 2014), when-similarity has lots of potentials to enhance social connectedness.

The idea of matching similar community members was also proposed in specific context such as health community where people may want to connect those who have a similar diagnosis (Carter, 2004; Hartzler & Pratt, 2011). Beyond matching people who share the same disease labels, researchers also recommended enabling people to search peers based on treatments, side effects, health knowledge, role, lifestyle, caregiving situations, and language style (Civan et al., 2009; Hartzler et al., 2016; Tixier & Lewkowicz, 2016). Specific dimensions such as attitudes toward recovery and temporal needs were mentioned in mental health community and peer programs (O'Leary et al., 2017). Some empirical programs (Andalibi & Flood, 2021) showed effectiveness of such similarity pair methods.

Intelligent Systems for CMC at Scale

In this chapter, I first review literature on mass interaction with regards to text-based and video-based communication, and live streaming as an integration of both modalities. I identified challenges in large-scale communication with goal of building online relationships. Then I review existing and potential solutions and dove deeply in one of them: the application of similarity effects. Though literature in communication and social psychology show contradictive results supporting or rejecting the similarity-attraction link, but there are a plenty of factors that may affect the results. I analyzed potential factors such as similarity dimensions, information delivery method, information exchanged in the interaction, type of relationship, and communication modality. Empirical experiments in CMC also showed opportunities of this design diagram. It is reasonable to assume similarity effects can be applied in large-scale live streaming if the variables are thoroughly considered.

In terms of pair or group assignment, some work assigned partners manually with the assistance of some tools such as spreadsheets (Andalibi & Flood, 2021). This method is qualitative and is more flexible. However, in the context of massive communication, the computation volume and time constraints make the tasks beyond human capacity. Therefore, algorithmic methods are widely adopted, such as counting the proportion of agreed items (Byrne et al., 1970) assigning weights to profile attributes (Wee & Lee, 2017), and computing cosine similarity (Cosley et al., 2003; Ludford et al., 2004). These algorithms are arbitrary and simple to use. With the advances of technologies, machine learning algorithms can be applied in such systems. For example, clustering algorithms can take user information as features and assign clusters in a short time.

When AI Determines Communicators

It is not hard to imagine, in the future, intelligent CMC systems equipped with machine learning algorithms, or AI, get involved in human communication. A computational agent can modify, augment, or generate messages, the so-called "AI-Mediated Communication (AI-MC)" (J. Hancock et al., 2020). Furthermore, an intelligent system can operate on behalf of users, not only process the message, but also construct communication dynamics by selecting, filtering, or recommending communicators. I refer such mediated communication as AI-embedded CMC. Although AI-MC and AI-embedded CMC were newly purposed concepts, tools and applications are easy to find. I list two as examples: people recommendation and conversation recommendation on social network sites (SNS).

People or Community Recommendation

People recommendation, or "social matching", is one of the most effective mechanisms to encourage social connections and grow networks (Guy, 2018). Since Terveen and McDonald (2005) proposed a research agenda of social matching, a large body of research work has been done to explore the opportunities of using computational methods to connect people. Besides introducing friend of a friend (Mcdonald, 2003), intelligent CMC systems can take advantage of a variety of user data, such as geolocation, tags, activity history, to recommend new friends (J. Chen et al., 2009), teammates, collaborators (Zheng et al., 2018; Zytko & Devreugd, 2019), and dating partners (Pizzato et al., 2012).

Besides recommending individual users, an intelligent system can also recommend communities, such as professional groups on LinkedIn (Sharma & Yan, 2013), location-based community (W. Zhang et al., 2013), or technical forums (Kumar et al., 2019).

Guy (2018) summarized three commonly used techniques to provide recommendations: 1) graph-based techniques take advantage of graph presentations of social network, such as follow relationships on Twitter and Instagram (Dahimene, Constantin, & Du Mouza, 2014; Guo et al., 2017); 2) interaction-based techniques consider the interactions among users such as liking, commenting, tagging (Fan et al., 2018); 3) content-based techniques take actual content into account, such as text and photos create by a user (Ding et al., 2013), items a user read or bookmarked (Guy et al., 2011). Even sensor data such as accelerometer and GPS signals have been used to infer lifestyle and used in recommendations (Z. Wang et al., 2015). Recent studies proposed hybrid methods, considering two or more features above to generate more personalized recommendations (Bertini et al., 2020; Karimpour et al., 2021).

Regarding the effects of these techniques, Chen et al. (2009) compared four recommendation algorithms in a within-subject survey study and a controlled between-subject field study. Results showed that social network-based algorithms received more positive feedback than content-based algorithms. While content-based algorithms performed stronger in discovering new friends.

Conversation Recommendation on Social Network Sites (SNS)

AI-embedded CMC systems can manipulate communication subjects by recommending conversations. Social network sites have a long history of tailoring algorithms used in content ranking (Oremus et al., 2021). Using algorithms to decide whose conversations shows at the top, or whose conversations can be seen by more people, shape the way people communicate. Authors of popular posts are more likely to have interactions with others, either via reactions, comments, or private messages. If a conversation with controversial content is mostly shown to users who tend to have anti-social behaviors, these is a good chance that aggressive arguments, harassment, and bullying will happen (Buckels et al., 2014; Shachaf & Hara, 2010). If such a conversation with trolling comments is classified to be “engaging” or “trending” by algorithm and distribute to more users, trolling behaviors may spread (J. Cheng et al., 2017). This ecosystem does not benefit the community building. If a conversation with creative content is mostly shown to users who share the same interests and like to contribute clever ideas, they may extend this conversation to an enjoyable and insightful one. If such content with valuable comments is classified as “engaging” and “popular” by algorithm and distribute to more users, people may find it a good community and want to have further interactions.

There are many methods to recommend conversations. Chen, Nairn, and Chi (2011) compared five algorithms that recommend conversations to Twitter users, and investigated the association between users' motivation and their preferences on recommendation algorithms. Through an online experiment, they found Twitter users with social purposes prefer tie-based algorithms than users with only information purposes. Zeng et al. (2022) proposed a conversation recommendation model that captures both global and local interactions. Global interactions reflect users' content and pragmatic preferences, represented by topic and discourse word clusters. Local interactions depict users' prior behaviors, including replying relations and chronological order of conversation.

This Dissertation: Designing Intelligent Interfaces to Facilitate Communication in Large-scale Live Streams

The contribution of this dissertation is to understand and facilitate mass interaction with the help of AI in the context of live streaming. Live streaming, an integration of video broadcasting and real-time text communication, is a relatively new communication approach and is becoming increasingly popular these days.

A human-centered design approach (Boy, 2011; Cooley, 2000) is applied in my series of work. The first step is to understand viewers' motivation, practices, and challenges when engaging with live streams. I posed the following research questions in Chapter 3:

> *What motivates people to watch live streaming on SNS? How do viewers engage with the streamer and other viewers in live streams on SNS? What challenges do viewers face when trying to engage with the streamer and other viewers and what causes them?*

Built upon the results from the interview study in Chapter 3 and a thorough literature review about the similarity effect as a potential solution, I proposed a similarity-based subgrouping method that divides viewers into subgroups based on similarity. This idea was evaluated in a within-subject controlled lab experiment using Wizard-of-Oz technique. The goals of the study are two-fold and the following questions are asked.

How does similarity-based subgrouping affect viewers' experience compared with no grouping and random grouping? What are the underlying relationships among factors that may affect viewers' watching experiences (e.g., ease of following the chat, conversational enjoyment, participation) and as well as perceptions of other audience members (e.g., interpersonal attraction)?

Based on the findings in the above experiment, and inspired by existing intelligent systems and explainable AI research (Abdul et al., 2018; Arrieta et al., 2020; Long & Magerko, 2020; T. Miller, 2019; Shneiderman, 2020), I designed Chatbuddies, a live stream chat system that applies clustering algorithms and generates subgroups for viewers in live streams (Chapter 5). Besides the clustering feature, Chatbuddies also visualize the group assignment and display viewer information to explain the algorithm. A within-subject online experiment is conducted to answer the following questions: evaluate the effects of Chatbuddies on viewers' use of the system and watching experience, compared with a control condition that generates subgroups randomly.

How do people use and perceive the visualization and subgrouping features of Chatbuddies? How does similarity-based grouping differ from random grouping in terms of viewers' watching experience and chatting behaviors?

In the next three chapters, I elaborate on my research studies and answer the questions above.

CHAPTER 3

UNDERSTANDING PEOPLE'S ENGAGEMENT WITH LIVE STREAMS ON SOCIAL NETWORK SITES

In this chapter, I aim to have an in-depth understanding of live stream viewers' motivations and behaviors on SNS, in order to identify challenges viewers face when trying to engage with others and to inform the future design of live streaming services.

Previous studies have explored viewers' motivations and practices on Twitch (Hamilton et al., 2014; Hilvert-Bruce et al., 2018). Researchers found social interaction, sense of community, meeting new people, entertainment, information seeking, and a lack of external support in real life are primary motivations. Lu et al. (2018) investigated Chinese users' motivations for watching live streams and listed relaxing, killing time, making friends, sharing opinions, finding a community, and gaining new knowledge as the most common motivations. As live streams became more embedded into SNS such as Facebook, Instagram, and YouTube and the viewership increased rapidly (Restream Team, 2021), there is an increasing need to understand viewers' motivations for watching and engaging live streams on platforms other than Twitch. Therefore, I list the first research question:

RQ1: What motivates people to watch live streaming on SNS?

Haimson and Tang (2017) investigated the attributes that make event viewing engaging by comparing viewers' experiences on Periscope, Facebook Live, and Snapchat Live Stories. They found that live streams especially offer interactivity and sociality for remote event viewing, compared with Live Stories on Snapchat. However, given those properties, current live stream services still have much space to improve. As Haimson and Tang (2017) pointed out, interactivity can drive or detract viewers' engagement, depending on comment content, comment volume, and the relationship between viewer and broadcaster. High comment volume decreases viewers' desire of watching a future event, which aligns with Hamilton et al.'s (2014) finding that the overwhelming text messages hinder meaningful interaction among the streamer and viewers. Given these existing problems, I hypnotize that more challenges may have not been discovered yet. To optimize current live stream tools in general to better support interactions among streamers and viewers in live streams on SNS, I need to develop a fine-grained understanding of viewers' practices of engagement with others to identify current challenges in live streams. Therefore, I list the second and third research questions:

RQ2: How do viewers engage with the streamer and other viewers in live streams on SNS?

RQ3: What challenges do viewers face when trying to engage with the streamer and other viewers and what causes them?

Methods

I conducted semi-structured interviews to gain an in-depth understanding of the viewers' motivations and ways of engagement when watching live streams. I also explored the challenges they are facing when trying to engage with the streamer and other viewers.

Participants

Participants were recruited from a large U.S. university through an online research recruiting system in Spring, 2018 and received extra credits as compensation. I interviewed 20 participants (10 = Male and 10 = Female) with an age range of 19-25. Participants' ethnicities were Caucasian (10), East Asian (6), African American (2), Bi-racial (1), and South Asian (1). Most participants were undergraduate students (18) and two participants were graduate students. Participants' native languages were English (16) and Chinese (4). All of them have smartphones and laptops. Participants watch live streams on a wide variety of platforms: Instagram (12), Facebook Live (9), YouTube Live (9), Periscope/Twitter (6), Twitch (3), Bilibili (2), and live-stream applications in other countries. Most of them reported watching live streams about twice or three times a week. Watching duration varied from 10 minutes to 2 hours.

Data Collection and Analysis

After a body of pilot interviews, I developed an interview protocol that asked participants questions about their motivations, experiences, and perspectives when engaging in live streams (the interview protocol is in Appendix A). I structured the interview into three phases. The first phase focused on the general experience of live streams. First, I gave participants the definition of live streaming in this study: the streamer uses a camera to broadcast real-life scenarios, such as traveling, outdoor activities, and talk shows. Viewers can interact with the streamer synchronously. Next, I asked participants about their general viewing of live streams (frequency, average watching duration, types of live streaming they watch and why, platforms they use, preferred and disliked features, and things doing when watching). To understand how they engage with the streamer and other viewers, I asked questions such as "Do you ever give reactions (e.g., hearts and emojis) to the streamer when watching on this platform?" "Do you ever chat with other audience members on this platform?" Based on the answer of the interviewees, I asked the reasons for doing or not doing it.

The second and third phases focused on the specific experiences to obtain detailed examples of participants' watching and engaging experiences in a certain context. I asked participants to describe one positive and one negative experience of watching live streams. Examples of questions include "Please think of and describe a live stream you watched that you thought the most engaging, interesting, and enjoyable." Participants were required to describe the experience in detail, such as when, where, and how they watched the live stream, the content of it, and what made them watch. I asked about their own experience or their observation of interactions between other viewers and the streamer. For the positive experience, I asked questions such as "What made it engaging, interesting, and enjoyable?" For the negative experience, I asked questions such as "What made this experience unpleasant?" Lastly, I encouraged the participants to add any thoughts about live stream watching that were not covered by the questions above.

Interviews lasted from 26 to 58 minutes (M=41). They were conducted in person and were audio-recorded, and then transcribed. Each transcript was numbered, and quotes are reported as Pn[gender] in the sections below.

All the transcripts were imported into Atlas.ti, a qualitative data analysis tool, and were broken into manageable pieces. I started with an initial open coding to generate labels from the interview data (Corbin & Strauss, 2012). Then I used axial coding to further categorize and compare the relationship between labels. In the end, selective coding was employed to generate concepts and themes and present the results.

Findings

In this section, I first explore the motivations for watching live streams on SNS. Next, I investigate how people engage with the streamer and other viewers. Lastly, I present the challenges viewers were facing when trying to engage with others in current live stream services.

Motivations for Watching Live Streams on SNS

When asked the type of live streaming they watched, participants listed a variety of categories such as political commentary, sports commentary, celebrities' talk shows, and friends' daily life. Most of them reported watching live videos out of interest in the topics or the unique content created by streamers. I also asked participants how they usually start to watch a live stream.

Most of the participants said they get notifications when browsing SNS; however, others choose from a live video list. Different from independent live streaming applications such as Twitch and Periscope, users encounter live videos without intuitively searching but "informally meet" when dwelling with daily social media. In this situation, participants have both individual and social motives for entertainment and information seeking.

Four themes related to motivation emerged from the data: enjoying interests, obtaining information, keeping updated, and connecting with others. I found these themes align with previous literature about the motivations of the video-based mixed media (de Oliveira et al., 2018). I incorporated the findings of their conclusions to expand the framework.

Individual Entertainment: Enjoy an Interest and Backstories

Users can watch live streams individually or along with others. When watching live streaming is an individual behavior, participants used it as a tool to benefit themselves, such as relaxing, killing time, learning new knowledge, and obtaining information.

Participants said they follow content creators or celebrities that they are interested in and would receive notifications when a live stream starts. They reported these live streams were mostly informal activities or conversations to share personal life and backstories, which arouse their great interest.

> *I knew he was moving from like another, like the backlogged video he had posted, and I wasn't watching anything else on YouTube. And then I saw that he was live streaming, and so I decided to watch and I wanted to see if he was going to be showing his new apartment yet because he should have been moving sometime that week. (P18, Female)*

Participants felt live streaming is a way they can know the personal lives of others. They saw being able to see celebrities' personal lives and backstories as a fun and enjoyable experience.

You don't really get to see what a famous person actually would do in their free time, like you find out a lot of gossip about them from the media and what they might have done that is so controversial, but then you find out what they do in their free time and they're just like you, they also have to follow a strict diet or... like Chance the Rapper has to hang out or doesn't have to but he chooses to hang out with his daughter in his free time. You would never know what a famous person's doing right now. [But] because of live stream you can know. (P14, Male)

Individual Information: Get Informed, From the Video and Comments

Some participants reported that they watch live streams to learn from the streamers or to learn new knowledge. They obtain information directly by watching a live video or by asking questions to the streamer. They also seek information from observation when the streamer informally presents their skills such as music performance and illustration.

Besides video, another approach to get information is from peer viewers' comments. They are interested in other viewers' real-time reactions which would be informative for them.

I'll just watch for a little bit, and I'm there mostly just to see what people are saying and their reactions, more than the content itself. (P8, Male)

I want to see what's the immediate reaction of other people. I do pay attention to other people's comments. I think that's another reason why I would watch live streams (...) Like, watching other people commenting on that is entertaining or educational enough for me. (P14, Male)

Social Entertainment: Say Hi, At the Same Time in Different Places

SNS are web-based services that allow individuals to construct profiles and create connections (Boyd & Ellison, 2007). The core difference between live streaming on SNS from previous live stream services is that they are embedded in the social media ecology. According to Joinson (2008) and previous studies on the motivations of using SNS, people consistently use it because they want to maintain social connections (e.g. connect with people who have not seen for a while, contact with remote friends), keep updated (e.g. see what friends are doing recently) and share identities (e.g. join groups and communicate with people sharing the same interest).

Some participants recalled being involved in their friends' live streams as the most engaging and enjoyable experience. For instance, an old friend streamed paragliding, a high school classmate showed her dorm in the university, or friends from a minority group publicly talked about their stories. Participants felt live streams provide an approach to connect with their existing social network. As P3 (Female) mentioned in the interview:

> *It's just daily things that my friends post, for instance, when they go out shopping, they sometimes, livestream. It's not like they have really meaningful content, but it's more a bonding thing because I just see each other and then perhaps leave a comment. And also my friends are mostly at other university campuses so I don't usually see our faces often. So it's mostly about seeing and checking on each other. (P3, Female)*

As most of the participants reported, they come into a live stream via browsing social media without planning in advance, which is an "informal" communication style similar to meeting someone in the hallway. P5 (Female) and P10 (Male) said they sent direct messages to the streamer after the stream ended.

Social Information: Keep Updated, See What Happens

As live streaming is embedded in SNS, it provides functionalities similar to text posts and images for status updates. Furthermore, it allows users to see others' concurrent activities and personal life which may not be revealed through other asynchronous mediums. Many participants felt curious when they saw friends' live streams as they were browsing SNS. Their purpose was to obtain information but not to interact with friends.

I had one friend who live-streamed his office hours. (...) I'm not even in physics classes and I watched it. He's good at explaining things. (P11, Male).

As I present above, these four themes are not mutually exclusive. Entertaining motives can be either individual or social, as well as information motives. I conclude the four motives and list examples in Table 2.

	Entertaining Motivation	**Informational Motivation**
Individual Motivation	Watch followed content creators' personal life	Learn knowledge from the streamer in video or comments sent by other viewers in live chat
Social Motivation	Connect with friends in real time	Keep updated and know what friends are doing recently

Table 2. Motivations for Watching Live Streams on SNS and Examples

Engaging with the Streamer and Other Viewers in Live Streams on SNS

I asked participants how they watch and engage with others such as their reactions and interactions with the streamer and other viewers. Participants reported various scenarios of engagement with others they experienced or observed in the live streams. I found several communication patterns and factors related to the engagement styles. The main theme that emerged from the data is the nature of the interaction relationship: some participants focused only on the streamer, while others also noticed the reactions from other audiences. I framed the results into an input-output model, in which input refers to participants' attention focus and output means their behaviors during watching the live streams.

Focusing on the Streamer

Live stream as mixed media combines broadcasting video and text chat channel. When watching a live video, the default interface will display live chat besides, underneath, or overlay the video, depending on the devices. Users have the option to display or hide the live chat.

Half of the participants (N=11) reported that they usually ignore the live chat and solely focus on the video when watching live streams. They revealed little interest in other viewers' opinions and reactions. Some of them usually hide the live chat or enter full-screen to get rid of live chat. Chat messages that keep scrolling at the side of or overlay the live video make it distracting.

Rarely, very rarely. I don't think people on Instagram generally like to read. I think I came here because I don't want to read things, I just want to watch things. (P3, female)

Although these participants focus on the video and streamer, they do give reactions such as emojis and hearts, which are anonymous and ephemeral. Interestingly, even though they are reacting, participants considered it as a passive way of watching and saw themselves as passive viewers.

Most of the time, I don't really comment. I'm more passive in terms of social media and the online community. But for hearts, if the person says something that resonates with me, I'll just give a heart. And the comments feel like they're more… I don't know. I don't think they were very useful for me. (P14, Male)

Using live streams as a background is another common activity because some live streams last for several hours and participants thought it was unnecessary to keep watching them. They had the live video on and do other activities on the side, such as doing housework, cooking, or having meals. Interestingly, participants did not find it difficult to understand the stream even without seeing the video or chat. Because usually, the streamer would read the comment they want to respond to.

> *I will sometimes peek and look—and actually, I can know what other people's question is because she will repeat (them). Normally they will repeat, and they will have some interactions, and I just listen. (P19, Female)*

When focusing on the streamer, video is the primary modality along with emojis or hearts as a lightweight component. Participants engaged with the streamer by watching the video as input and sending reactions as output.

Engaging with Both the Streamer and Other Viewers

Live streaming consists of both live video and text chat. Some platforms also include gifting and donation features. Users have multiple approaches to communicating with the streamer and other audiences. In this situation, they take video, text chat, donations, and other modalities as input. They read the comments and noticed the gift donated by other viewers. They thought these modalities help to construct the live video content.

> *(...) Obviously, the comments are a big factor because it's the viewers' response, and that's such a big part of live streaming. It's not just the speaker. They'd need to have the comments because they'd be nothing without the viewers. It's more like an engagement between viewers and speaker, and it wouldn't just be solely the speaker. (P15, Female)*

Viewers have various ways to engage with the streamer and other viewers when they communicate via multiple modalities. I listed some common engagement behaviors below.

Respond to the streamer. Some participants (N=9) reported sending text messages to the streamer occasionally, such as asking questions, responding to the streamer's questions, showing compliments, or commenting for incentives.

A lot of times I'll do reaction-type things. Like they'll ask for comments and then you try to comment to see if they can recognize you because then they might end up following you and then you have a famous person following you which would be cool. (P12, Male)

Speak out consciously. Some participants (N=8) talked about their experience engaging with both the streamer and other viewers, but mostly in a conscious way. They comment with a specific purpose, such as revealing opinions, arguing with others, expressing emotions, or connecting with the community. As P1 said below, she argued with other viewers to support the streamer.

Sometimes when people say "oh it is boring" or something, I'll become a little bit angry, and will try to stand on the host's side and just... a little bit argument. Just to say something like "oh this makes me happy, and if you don't like..." or something like that. Because I don't want the host to feel frustrated because of the comments. (P1, Female)

Join a personal conversation. Participants consciously choose different ways to interact. Most of them tend to participate more in friends' live streams including leaving a comment or have a conversation instead of just sending hearts or emojis. They felt more comfortable speaking out in an environment where the audience includes people they know.

I would definitely leave a comment if he or she is like a close friend of mine. Like "I'm watching it", or just have a conversation like "how's it going?". (P3, Female)

Usually when it's a friend, I'm more likely to comment because I feel more comfortable expressing my opinions, but on famous people's or bloggers' live streams, I'll sometimes do hearts, but I won't frequently comment. (P15, Female)

Some participants described the experience of interacting with friends by sending text messages in the live chat and getting verbal responses in real time. They thought it was similar to a normal conversation and felt intimate.

If you're watching a livestream of one of your friends, it definitely feels like a more intimate experience. Like you can directly ask them a question and they'll see it right away and then they can answer it verbally. (P16, Female)

If it's someone famous, they already know they're going to have audiences when they go on. But there were only four or five people in my friend's stream. I talked back and forth like a normal conversation. (P13, Male)

Self-expressing. Few participants (N=2) reported sending spam and random messages to live chat frequently. They used the chat feature to express their reactions to the live video. They described the chat feature as fun and cool as their messages fly away quickly. In this situation, there were little interactions with other audiences or the streamer.

I think that's because Instagram live video goes away, you feel your comments are just going into the abyss. (...) I'll comment on there because it gets lost in the stream of comments. If you're on Instagram and you comment, it's gone. There are just so many messages depending on whom you're watching that it just goes away. So, I'm okay with doing that. (P11, Male)

Lurking. Lurking is another common behavior reported by participants (N=10) (Preece et al., 2004). Although they felt enjoyable when watching the video and reading the comments, they did not have an interest to participate in the chat, especially in a large-scale live stream, where they did not have personal connections with either the streamer or other audiences.

I don't have the obligation to talk to that person, because it could be one-person live streaming and there could be a thousand people on the other watching the live stream. (...) I don't feel that I have the social pressure —"pressure"—to talk to that person. (P14, Male)

In addition, participants thought even if they keep silent, there would be other viewers' comments coming in and individuals would have little contribution to the stream.

As Preece et al. (2004) pointed out in their study of lurking, there are a host of reasons for people to lurk, but lurkers are not free riders. Even though they would like to, participants in my studies revealed the challenges they were facing when trying to engage, which I will discuss in the next section.

There are some other engagement approaches on some live stream platforms. For example, viewers can send gifts or financial support to the streamer (Lu, Xia, et al., 2018; Wohn et al., 2018), or join the stream and video chat with the streamer. But in my study, no participants reported having this experience.

Interestingly, these two engagement styles, focusing on the streamer and noticing both the streamer and other viewers, are not static. Even though usually interviewees had preferences, they reported that sometimes they shift attention dynamically as the content of live video and engagement behaviors changed. For example, P3 (Female) revealed that she usually did not read other viewers' comments, but when she wanted to ask questions to the streamer, she often took a look at others' questions first. Another participant talked about watching a famous music producer's live video. The live stream became less engaging when the broadcaster left the camera to get some food, this participant started to read live messages and chat with other viewers. After the streamer came back, he focused on the live video again.

It kind of goes away from the live stream and into live chatting. (...) But sometimes just because of whatever the person is doing, it gets kind of boring or they're not trying to be the focus of attention like they're just eating. The live chat becomes its own little community. (P11, Male)

Challenges of Engagement

I identified the following challenges viewers face when trying to engage with others in live streams: distraction, lack of direct communication with the streamer, lack of coherent and meaningful conversations, lack of a friendly environment, and privacy. I present the potential causes of these challenges below.

Distraction

For viewers who tend to focus on the video, either emojis or chat messages could be distracting and attention-demanding. Most live streaming services on SNS display live chat by default. Some platforms allow users to hide the comments or use full-screen to avoid the text chat, but others do not, which makes it difficult for viewers to concentrate.

It's just distracting frankly. Like I don't have anything to say in terms of the comment section so I'm just trying to watch the content, I don't really need to see everyone else's opinion. (P6, Male)

For participants who would like to engage with both the streamer and other viewers, emojis, text chat, and other components brought them a mixed feeling. On one hand, knowing others were doing the same thing and seeing their real-time reactions made them feel engaged. They found it interesting to see others' comments and get a sense of connection. On the other hand, they found it became distracting when they wanted to focus on the video. They need to shift attention back and forth, which is mentally demanding. This finding echoed previous research on the effects of chatting when co-watching videos (Weisz et al., 2007).

It's sometimes annoying because for Facebook, if I press likes and hearts and stuff, all the icons will float around and people will get annoyed at that. I also get annoyed by that because I get distracted, but it does kind of add to the feeling like "oh, I'm watching a live video." So it's kind of like a mixed feeling of good and bad, but if it's a video I really want to focus on, that does distract me. (P3, Female)

It definitely does get distracting after a while. But it is also very easy, like an easy way to see the general feelings about the live stream. So it has its drawbacks but also like its positives. (P8, Male)

Balancing the distraction from the live stream components and the desire to learn about periphery information is a challenge for viewers.

Lack of Direct Communication with the Streamer

In a large live stream, participants felt it is unlikely that the streamer will see their comments because he/she is reading a huge number of messages. The rapid scrolling speed makes it nearly impossible for the streamer to go through each comment.

I feel my comment wouldn't really do anything because they're getting so many. I feel there's such a low likelihood they would even read it because they're talking and people are also commenting at the same time. They're not really doing both unless they take a second to start reading all the comments. (P15, Female)

Participants expected their voices could be heard when interacting with the streamer. Although some platforms provide features for viewers to be recognized by the streamer, such as donating and being a guest, participants felt uncomfortable and rarely use them. They prefer a more natural way of direct communication.

The one-to-many communication pattern is not preferred by some participants. Although the text channel can be used to communicate with the streamer, all the messages would also be seen by other viewers. P5 (Female) revealed that she disliked this unbalanced communication.

> *I still do like a one-on-one conversation. I feel it's a little bit of an unbalanced communication on the live stream because you can see them and everything they're doing. Meanwhile, they're just getting the messages. I don't really like that dynamic very much. (P5, Female)*

Lack of Coherent and Meaningful Conversations

Another challenge is that participants found it hard to read or participate in coherent and meaningful conversations. As Hamilton et al. (2014) pointed out in the early Twitch study, conversations broke down due to information overload. When there were a large number of comments scrolling, participants complained that it was difficult for them to find the one they were interested in. The single message fell into the abyss and required much effort to find again.

> *If it's someone famous then you have to scroll really fast to find it again. Or it just gets lost in the sea of people just commenting so fast. If you see it one time, a lot of times you just have to remember that someone commented because it's impossible to find it again. (P12, Male)*

Moreover, the incoherence hindered participants to join the conversation. As the text chat kept moving at a fast pace, it required them to react quickly to locate the message and type back.

It’s hard to reply when there’s just so much going on because you’ll see a comment and *you’ll want to reply to it but then it’ll be gone as soon as you even think of something to say, so that’s one of the main reasons I didn’t reply. (P8, Male)*

Conversations can be meaningful in terms of the content and the subject participants talked to. Participants who were seeking information were interested in other viewers’ comments and they wanted to read more informational and educational conversations instead of mundane comments or messages that were unrelated to the live video. They thought meaningful conversations made watching experiences more engaging.

I think when people have more meaningful conversations instead of just saying random things or things that are not really in relation to the live stream, I think people have the real conversations they are actually discussing. [The conversation] makes it more engaging. (P8, Male)

Participants also felt there was no point to engage with strangers in live streams. They thought they were random people who happened to watch the same live video together. On the contrary, for participants who were seeking social connections, engaging with friends was meaningful even though they were not doing any activities with a particular purpose. Participants said they appreciated the chance to talk with friends in real time, either when their friend was the streamer or viewers. Because they knew their friends would see it and it was meaningful for them.

I don't really care that much to express an opinion or to comment on a famous blogger’s post. (...) As opposed to a friend, I know that they don't have millions of people on their live stream, and I’m also close with them, so I know that if they see my comment, they’re going to answer it or appreciate it and it's going to mean something. (P15, Female)

For some participants, "friends" does not have to be individuals with personal connection in real life. Instead, they can be people with whom they have connections in online communities. As Hamilton et al. (2014) presented, streamers and viewers build close-knit communities on Twitch by actively interacting regularly. Active users usually have intimate friendships.

But for viewers who do not participate regularly in a certain streamer's live stream, it is unlikely to expect they will directly find social connections with other viewers. Even if friends show up in live streams in large-scale live streams, most current systems do not notify viewers. It is difficult for them to find or join socially meaningful conversations.

Lack of Friendly Communication Environment

Similar to challenges in many online environments, anti-social behaviors and inappropriate utterances bring negative experiences and hinder users' intention to participation. Participants reported seeing aggressive, opinionated, and polarized comments. They said they would rather not see the live chat if the messages are toxic.

> *If there was no chat, I think I'd be less distracted and it would make me less angry than just watching it. I mean I don't love this president, but I think that this chat just got me extra hyped about everything he stands for, where if I was just watching him speak, which is kind of what I wanted to do, I think I would've been calmer [sic]. I mean I wouldn't have been calm, I don't like him, but like I would have been calmer [sic] not seeing thousands of people too. (P2, Male)*

Participants thought it was frustrating when they have to see those inappropriate messages, such as trolling. It detracted their intention to stay in current live streams, even watch live streams in the future.

I usually don't engage, especially when I see people are just trolling, cause it's not even worth it. So that does make it I guess frustrating because you have to see what they're saying and it makes you less likely to maybe use that live stream again or to stay on the live stream. (P8, Male)

Privacy

Privacy is a huge issue in social media technologies (Ziegele & Quiring, 2011). On SNS, people are aware of their identity and personal information and prefer to hide their watching history.

Because on YouTube, they don't have Danmaku, and it's not anonymous anymore. If I comment on the video, it will just show my name [and] they can fight me through Facebook, so it's not anonymous. Also when I watch the video, I'm just trying to release my pressure. (...). I don't want anybody to know that. (P1, Female)

This kind of concern also exists in friends' live streaming. Even though the streamer is a personal friend, if the relationship with other audiences is not intimate, some participants preferred to keep the conversation private between them and the streamer.

Because I wouldn't want someone else to see my conversation in the stream. If I'm gonna text the person, I'd rather do it on the side... not everyone's friends watching it. (..) It's good to not be seen all the time... I don't know how close the mutual friends are and whether I want them to see what I'm saying." (P10, Male)

Discussion

In this study, I explored the motivation, practices, and challenges of watching and engaging live streams on SNS in a fine-grained manner. I summarized the motivations from an individual or social, entertaining, or informational perspective. Participants' engagement styles and communication approaches are affected by their motivations, relationship with the streamer and other viewers, as well as current communication environment. I framed the findings into an input-output framework and show them in Table 3 below. There are two types of roles, loyal followers who focus on the streamer, and community players who actively engage with other viewers. Participants usually have a preference of engagement style, but they can also switch between the two depending on the live stream category and content.

Input	**Output**	**Challenges**
Focus on the streamer (Loyal follower)	React Respond to the streamer	Distraction Lack of direct communication
Not only watch video, also read the live chat (Community player)	Join a personal conversation Speak out consciously Self-express	Lack of coherent and meaningful conversation Lack of friendly communication environment Privacy Concerns

Table 3. Engagement Styles and Challenges

As Hamilton et al. (2014) and Haimson and Tang (2017) suggested in their papers, when the group size grows larger, subgrouping is an effective approach to facilitate interactions and meaningful conversation. To solve the challenge of information overload in live streaming chat, Miller et al. (2017) designed a conversational circle algorithm that randomly assigns an audience into a neighborhood and uses an upvoting function to filter the messages the streamer can see. The evaluation shows a positive effect with respect to the easiness to use and understanding, supporting the community, and highlighting important content. Their design indicates that subgrouping and upvoting strategies are effective in dealing with information overload and incoherence in conversations.

Future design could move forward on subgrouping strategies and go beyond random grouping. Constructing subgroups using various strategies could have the following benefits: (1)

It can help generate meaningful conversation content. Participants found it frustrating to keep seeing many mundane and irrelevant comments. If the system generates subgroups by people who are similar or dissimilar, it will be more likely to see comments that are useful for them. For example, P3 (Female) mentioned she would be more willing to read comments from viewers of similar age when watching makeup live streams. P20 (Male) said he had a nice conversation with a viewer who lived near his hometown when watching a fishing live stream; (2) It can help viewers find valuable social connections. Based on my findings, interviewees are more willing to interact with someone they know or are from the same network. Weisz et al. (2017) found that watching videos and chatting with friends in real time increases the feeling of closeness. It is possible that the future design could incorporate subgroups according to users' existing social connections. Reading comments or chatting with a friend or a potential friend should be more engaging and meaningful than interacting with a group of random people. (3) It can help moderation and create a friendly communication environment. Douglas and McGarty (2001) found that identifiable group members tend to act in a more group-normative manner. Within a smaller-sized and consistent group, members might be more self-regulated and use more acceptable languages.

I also found users have a demand for subgrouping even when the audience size is not large. Participants have privacy concerns when other viewers do not share intimate connections. They prefer private conversation not to be seen by others. In large-scale live streams, they tend to lurk and keep silent; while in friends' live streams, they may text friends via other communication tools. Generating subgroups among mutual friends and providing opportunities for private interaction will be beneficial for developing intimate connections.

The primary limitation of this study is the small sample size and lack of diversity. I only interviewed young adults in a university setting because they are a convenient sample and they are active on social media and live streaming services (Lu, Xia, et al., 2018). This population may not present other user groups in the live stream industry.

In the next chapter, I present an experiment that evaluates the effects of subgrouping on viewers' perspectives on livestreams.

CHAPTER 4

EVALUATING THE EFFECTS OF SIMILARITY-BASED SUBGROUPING ON AUDIENCE EXPERIENCE IN LIVE STREAMING - A LAB EXPERIMENT

The literature reviewed in Chapter 2 indicates that creating subgroups based on similarity can be an effective way to affect participants' perceptions and behaviors in online communications. The concept of community has been discussed in previous work on live streaming, especially when the streamer is devoted to constructing a community (Hamilton et al., 2014; J. Li et al., 2019; Lu, Xia, et al., 2018). Live streaming is different from other online communities in the following aspects. First, the concept of community is transitory in some live streams, such as in celebrity streams and large-scale events where the broadcaster is not dedicated to building an online community. Second, live streaming follows the framework of a one-to-many system where the streamer, the center of the media, communicates with a large audience at the same time. Audiences have a variety of communication preferences. Some of them send direct messages to the streamer; some exchange ideas with other audience members; others prefer to "lurk" and read others' comments instead of actively participating in the discussion.

I lay out the hypothesis model and research questions (Figure 3). As stated above, the goals of this study are two-fold: (a) to propose and evaluate a similarity-based subgrouping strategy in live streaming and compare it with two baseline conditions (b) to investigate the underlining relationships among factors that relate to viewers' watching experience and perception of others. I am interested in two outcomes of livestream watching experience: liking of the stream and interpersonal attraction. The former serves as a measurement of individual users. Liking of the video and viewership may contribute to the popularity of the video (Pinto et al., 2013), which is an important measure for livestream platforms and streamers. The latter serves as a measurement of the group. Social psychology literature indicates that a community survives longer and has more group cohesiveness when online users have bond-based commitment such that they build connections with other people in the community (Kraut & Resnick, 2012). Hamilton et al. (2014) found that live stream users often cluster as a community around the streamer they follow, and some streamers devote themselves to creating such a community. Therefore, I explore whether the proposed grouping strategy will have effects on these two outcomes.

In addition, I consider other factors that are relevant to the watching experience, including ease of following the chat conversational enjoyment, participation, mental workload, and mental engagement. These factors have been used in previous studies (Haimson & Tang, 2017; M. K. Miller et al., 2017), but the underlying relationships among them are unclear. Little literature investigated the relationship between subgrouping methods and these factors, let alone whether there are any mediation effects. In the later part of this section, I present the hypothesis based on literature, theories, or logical deductions; I raise research questions for factors if there is little evidence to generate a hypothesis.

Many people complained that the huge volume of chat makes it hard to follow conversations and makes the experience less meaningful (Haimson & Tang, 2017). My goal is to enhance the user experience of the chat and help people find content relevant to themselves. Both chat volume and chat content may affect viewers' perceptions of chat and their participation. In Miller et al. (2017)'s study of the "neighborhood" technique, for example, participants reported the chat was easier to respond to and more balanced in terms of volume. Subdividing viewers into groups manages the chat volume in each group and should make it easier to follow the chat.

H1: Following livestream chat conversations within a subgroup will be easier than without a subgroup.

Conversational enjoyment is a psychological measurement that captures how people like the conversations, which usually happen after people read or participate in the chats. If people cannot catch up with the chats, it is less likely they will get involved or enjoy the conversation. Therefore, I assume that increased ease of following the chat may then enhance conversational enjoyment.

H2: Viewers' conversational enjoyment will be positively associated with the ease of following conversations.

The ease of following the chat may motivate viewers to participate in the conversation. If people find the chat easier to read, they may want to contribute to the conversation. But other factors also affect online participation, such as knowledge about the topic, attitudes towards online activities, design (McInnis et al., 2016), and personality traits (Craker & March, 2016). It is not clear how the ease of following the chat will contribute to participation. Therefore, I propose the following research question:

RQ1: How does ease of following the chat affect viewers' participation in live streams?

Decades of social psychology research indicate that social interaction increases liking (Hogg, 1993). The conversational experience may affect the liking of the live video itself and the liking of other group members. An early study conducted by Weisz et al. (2007) showed that people like each other more when watching TV and chatting with others versus not chatting. Going beyond chat functionality, I hypothesize that conversational enjoyment will affect people's liking of others, such that the more an individual enjoys the chat, the more they will like other viewers who participate in the chat. I propose the following hypotheses:

H3: Viewers' liking of the live stream will be positively associated with conversational enjoyment.

H4: Viewers' perception of other audience members will be positively associated with conversational enjoyment.

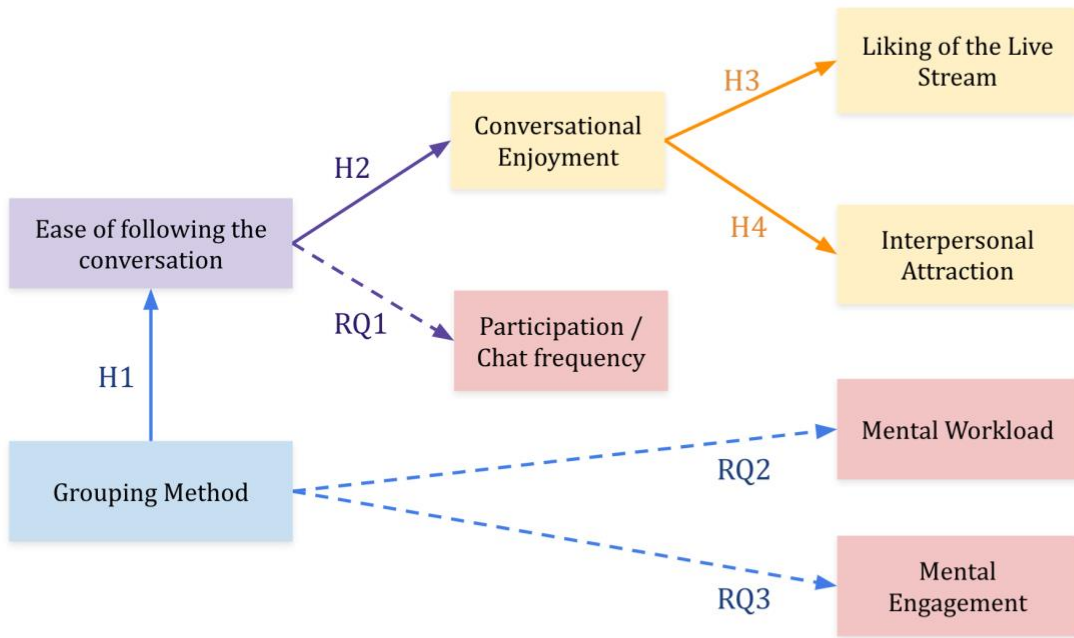


Figure 3. Hypothesis Model

Interpersonal attractiveness and liking of the live stream were not hypothesized as predictors of conversational enjoyment because I did not tell participants how the subgroups are generated. Therefore, participants should construct their impressions of other audience members from the conversations instead of vice versa. Similarly, liking of the live stream emphasizes the enjoyment of watching instead of the initial liking of the stream topic.

According to resource theory (Kahneman, 1973; Lang, 2000), people can allocate mental resources when multitasking in order the achieve the best performance. When the demand exceeds the available resources, the performance will decrease. Reading a huge number of messages requires more cognitive resources than a few messages. Therefore, increasing the ease of following the chat may reduce viewers' mental workload. Exposing audience members to messages that are relevant to them may increase their mental engagement. But live video and text chat are experienced as a whole, it is not clear when participants pay more attention to the live video versus the chat. Therefore, I raise the following research questions:

RQ2: How does subgrouping based on similarity affect audiences' mental workload when interacting with live streams?

RQ3: How does subgrouping based on similarity affect audiences' mental engagement when interacting with live streams?

Methods

In this study, I examined the effects of similarity-based subgrouping on ease of following a live stream chat, conversational enjoyment, liking of the live stream, and interpersonal attraction. I designed a within-subject lab experiment in which participants watched and engaged with three simulated live streams, each in a different grouping condition. To ensure that the livestream video content and chat messages were consistent across all participants, archived live videos and chats were used to simulate a "live" experience. Participants were "grouped" with chatters who had sent messages when the stream was live. Participants were led to believe that they were chatting with these people in real time.

Similarity has many dimensions. A number of these dimensions are implicit in online environments, such as age, gender, and ethnicity. This kind of information was not accessible to "other audience members" because they were audiences of the archived live videos but not participants in this study. Therefore, I chose a more explicit dimension, the similarity of interests, which can be inferred from the chat messages. For example, in a traveling stream, some people are interested in the personal experiences of the streamer, some are interested in the landscape and buildings the streamer is visiting, and others are interested in the culture and history. Interests were inferred from the chat messages they sent following the logic that a person who is interested in a topic will send messages about that topic.

To generalize the hypotheses, I selected three types of live streams: travel (outdoor activity), grocery shopping (indoor activity), and knowledge sharing (webcam). I designed a 3 (grouping algorithm) by 3 (type of stream) within-subjects experiment with grouping and stream type counterbalanced across participants.

Participants

Participants were recruited from an online recruitment system at a large U.S. university. They received extra credit or $10 as compensation. Three participants were excluded from the analysis because they disclosed that they recognized the video was pre-recorded and the chats were not sent by real users. I report results from the rest 36 participants (14 = Male, 21 = Female, 1 = Prefer not to disclose) with an age range of 19-23. Participants' self-reported ethnicities were White (19), Black (3), Asian (13), and Other (1). Among the participants, 16.7% (6/36) reported watching live streams more than once a week, 22.2% (8/36) reported watching 2-3 times a week, 27.8% (10/36) reported watching once a month, and 33.3% (12/36) reported never watching a live stream. Slightly more than half (56.6%, 20/36) reported that they had sent chat messages in a live stream.

Materials

I implemented a real-time chat Chrome extension that can be embedded in YouTube. The interface is similar to the original chat window in YouTube Live (Figure 4). It was developed using the React.js framework and Firebase. Users can sign up and join the chat channel. The functionalities are the same as a normal chat room, the only difference is that an admin (the experimenter in this study) can load pre-curated messages into the chatroom.

“Live” Videos.

In this study, I chose three types of live streams: travel (outdoor activity), grocery shopping (indoor activity), and knowledge sharing (webcam). In the travel stream, a male streamer showed viewers around a small town in Europe and talked about his travel experience. In the grocery shopping stream, a female and a male streamer selected food in an Asian grocery store and talked about the food they saw. In the knowledge sharing stream, a male streamer answered questions from audience members about psychology and health. All three streamers have more than 1,000 subscribers on Twitch or Periscope. Their streams are publicly available. The experimenter downloaded and trimmed the streams after the streamers granted permission. A 20-minute snippet with no inappropriate or offending content was selected from each stream for use in the study.

Figure 4. Screenshot of Grocery Shopping Live Stream (Random Grouping Condition)

“Live” Chat and Group Generation.

The original chat logs of the three live streams were transcribed. Two researchers with experience in live stream chat annotated the chat messages into four categories: irrelevant content, target topic, other topics, and personal comments about the streamer. The codes were generated after the two researchers watched the selected live streams and discussed the experiment purpose. Each of them coded the entire chat dataset and 78% of the codes matched. Disagreements were resolved through a post-annotation discussion session. Target topics for the three videos were scenery (travel), food (grocery), and mental health (knowledge). Other topics in the videos include history (travel), culture (travel), travel tips (travel), cooking (grocery), diet (grocery and knowledge), and weather (knowledge).

Messages in the three conditions were generated as follows. In the no grouping condition, I used the original messages for the live stream. In the random grouping condition, I randomly selected a subgroup of chatters and made the group size similar to the similarity-based grouping condition. In the similarity-based grouping condition, I computed the ratio of each message category for all the messages a viewer sent to the channel. For example, a viewer in the grocery shopping live stream sent 10 messages in total, 50% were about food, 30% were about cooking techniques, 10% were emoticons, and 10% were personal comments about the streamer. I selected a subgroup of chatters who sent more than 60% of messages on target topics. This percentage was chosen after several pilot experiments that the chats have a moderate volume and readable speed (see Table 4). Chat volume can greatly affect watching experience (Haimson & Tang, 2017), therefore I tried to control this variable between random and similar conditions. As a result, viewers in subgroups that were created based on similarity were those who talk mostly about the target topics.

Condition	Travel		Grocery		Knowledge	
	Chatters	Speed	Chatters	Speed	Chatters	Speed
No Grouping (NG)	227	1.16	358	1.64	96	0.56
Random Grouping (RG)	42	0.24	48	0.21	32	0.2
Similarity-based Grouping (SG)	41	0.22	49	0.22	31	0.18

Note: chat speed = number of messages per second

Table 4. Number of Chatters and Chat Speed in Each Condition

Procedure

First, I introduced the study purpose and procedure. After participants installed the Chrome extension, they were told they would have three trials of study activities and a final survey. In each trial, the experimenter told the participants that she just found a live stream and shared the video link with them. The participants were told they should pay attention to a certain aspect of the live stream, and they need to answer some questions afterward. For example, they were told that for travel and outdoor live streams, they were instructed to pay attention to the scenery, landscape, and buildings. I primed participants in this way to manipulate their interests in the live streams. Priming is a common and effective approach to implicitly influence participants' cognition in behavioral experiments (Cameron et al., 2012). Although participants may not be sincerely interested in the subtopic, priming is an effective approach to orient participants' attention to target topics at least in the experiment.

Participants were told they should feel free to interact with the streamer and other viewers while watching but it was not mandatory. Each watching trial lasted for around 15 minutes. After each trial, participants completed a survey about their experience during that trial. Items in the post-trial survey are described in Appendix B.

Across the entire trial, each participant experienced all three kinds of grouping (no grouping, random grouping, and similarity-based grouping) and all three types of live streams (travel, grocery shopping, and knowledge sharing). The combination of grouping algorithms and livestream types was counterbalanced, such that each participant experienced each grouping algorithm once and each live stream type once. After three trials were completed, they answered a final survey. Questions in the final survey include live stream watching habits, such as how often they watch live streams, how often they chat in the live streams, and the functionalities of live chats in their mind. I also asked manipulation check questions, such as asking participants to describe the live streams they watched just now, how much attention they pay to the target topics (scenery, food, and mental health), and how much attention they believe other viewers paid attention these topics in the three live streams. In the end, participants answered demographic questions (see Appendix B). The study procedure is shown in Figure 5.

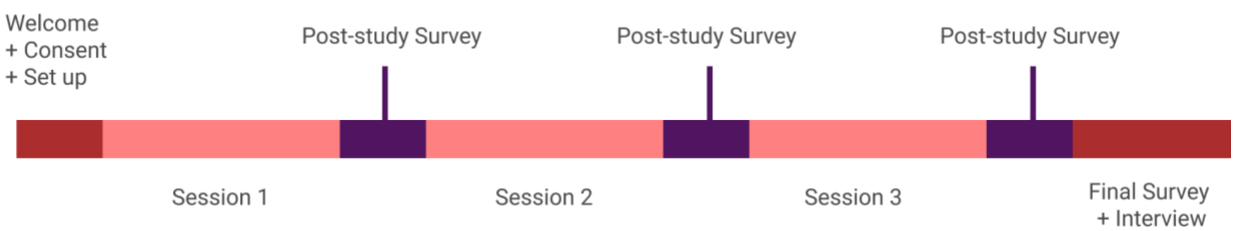


Figure 5. Study Procedure

I did not tell the participants whether audiences in the subgroup had shared interests with them or not, because I want to let the participants form an impression of other viewers from the chats instead of pre-gained information. Previous research has shown that perceived similarity can predict interpersonal attraction in the short-term interaction (Montoya et al., 2008), but the perceived similarity is difficult to manipulate given an existing set of messages.

In each watching trial, the experimenter broadcasted the archived live video through three different YouTube accounts. Videos were broadcasted using StreamLabs on a Windows 10 desktop computer. Once the participants opened the live stream, the experimenter loaded the curated chat messages to the chatroom using the Chrome extension. From the perspective of a viewer, it looked exactly like a real live stream. To simulate the interactions and control the impact of reciprocity, the experimenter replied to participants who sent chat messages at most twice. The replied messages addressed participants by having "@<username>" in the message body. Message content was selected from a prepared pool of messages that shows greetings and agreement, such as “hi”, “you’re right”, and “I heard that too.”

After participants completed the final survey, I interviewed them about their watching experiences and their impressions of other viewers. Questions included: “How much attention did you pay to the chat?”, “What do you think of the video/chat/other viewers of the first/second/third live stream?”, “How would you compare the video/chat/other viewers in the three live streams?”, and “Is there anything that you were not satisfied with?” The interviews were audio-recorded and transcribed to text. I debriefed participants on the use of prerecorded video after the study ended. Three participants revealed that they did not believe the video was live or the chats were sent by real people. Therefore, I excluded their responses.

Measurement

Both survey measures and system log data were collected. I am interested in two ultimate outcomes: Liking of the stream and Interpersonal Attraction, which may indicate viewers’ future behaviors in the live stream channel and the channel community. I want to explore the underlying relationship between other factors that were introduced in previous studies. As I demonstrate above, I construct two mediator variables, Ease of Following the Chat and Conversational Enjoyment. I am also interested in viewers’ participation, their mental workload, and engagement.

Ease of Following the Chat. I adopted the items used by Miller et al. (2017) to measure the volume of incoming messages and ease of use: “I was able to read all the text chat messages”, “I could reply to messages easily”, and “I felt the balance between too few and too many messages.” Participants indicated on a 7-point scale ranging from “Strongly Disagree” to “Strongly Agree”. Factor Analysis with Varimax Rotation indicated that all items fell on a single dimension accounting for 80.96 % of the variance, so I averaged these three items to obtain the measurement of ease of following the chat (Cronbach's alpha = .91).

Conversational Enjoyment. I adopted four items based on Ryan and Deci (2000) about interest and enjoyment. Questions included: "I had so much fun participating in this conversation when watching the live stream", "I thought the conversation was enjoyable", "I enjoyed the conversation with others when watching this live stream", and "The conversation did NOT hold my attention at all (reserved)". Participants indicated on a 7-point scale ranging from "Strongly Disagree" to "Strongly Agree". Factor Analysis with Varimax Rotation indicated that all items fell on a single dimension accounting for 71.58 % of the variance, so I averaged them to construct this measurement (Cronbach's alpha =.86).

Participation. I stored the messages participants sent during each live stream. I measured their participation by counting the number of messages they sent in each live stream. As the histogram in Figure 6 shows, participants' chat frequency was not normally distributed. Therefore, I converted it to a binary variable. Participants who sent at least one message were classified as 'participate', and those who didn't send any message as 'not'. The percentage of participation in three Grouping Methods conditions was 41.67% (15/36) in NG, 52.78% (19/36) in RG, and 55.56% (20/36) in SG.

Liking of the Live Stream. Two items based on Weisz et al. (2007) ("I had fun watching the live stream" and "I enjoyed watching the live stream") were averaged to get the measurement "video likeness" (Cronbach's alpha = .95). Participants indicated on a 7-point scale ranging from "Strongly Disagree" to "Strongly Agree".

Interpersonal Attraction. I measured participants' perceptions of others on Interpersonal Attraction, Sense of Community, Shared Interest, and Perception of Similarity. Interpersonal

Attraction was measured by four items based on Ren et al. (2012): “I felt close with them when watching this live stream”, “I would like to be friends with them”, “I am interested in learning more about them”, and “I like these people”. The other four questions measured Sense of Community (M. K. Miller et al., 2017), Shared Interest, and Perceived Similarity respectively: “I felt there was a strong sense of community when viewing this live stream” and “I think other viewers and I have shared interests”, and “Other viewers are similar to me”. Participants indicated on a 7-point scale ranging from “Strongly Disagree” to “Strongly Agree”. Factor Analysis with Varimax Rotation indicated that all seven items were on a single dimension. Therefore I averaged them to create the measure of interpersonal attraction (Cronbach’s alpha =.93).

Mental Workload. Participants were asked a series of questions about their mental demand, temporal demand, performance, the effort required, and frustration level (NASA TLX) (Hart & Staveland, 1988). Participants indicated on a 7-point scale ranging from “Very Low” to “Very High”. Factor Analysis with Varimax Rotation showed that performance and frustration loaded on a separate factor than the other four questions, which loaded on a single dimension. Therefore, I omitted performance and frustration and averaged the other three questions to create the measure of workload (Cronbach’s alpha = .75).

Mental Engagement. I used the definition developed by O’Brien & Toms (2008) that engagement is a quality of user experience constructed by four attributes: focused attention, endurability, novelty, and felt involvement (O’Brien & Toms, 2010). Example questions include: “I forgot about my immediate surroundings while viewing”, “My watching experience was rewarding”, “I felt interested in my watching task”, and “I was so involved in the watching task that lost track of time”. Participants indicated on a 7-point scale ranging from “Strongly Disagree” to “Strongly Agree”. Analysis with Varimax Rotation indicated that all items fell on a single dimension, so I averaged them to construct the measurement of mental engagement (Cronbach's alpha = .93).

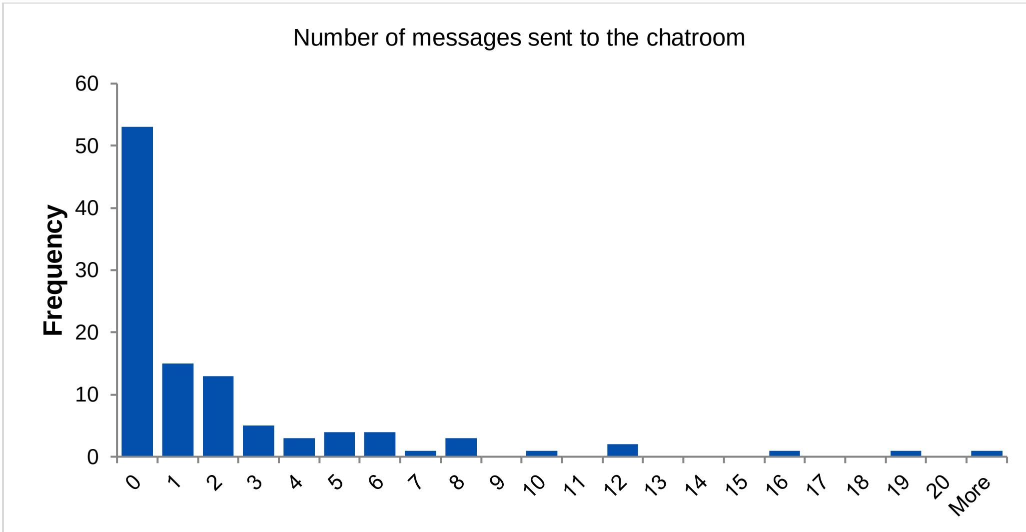


Figure 6: Number of Messages Sent to the Chatroom (N=108)

Data Analysis

The quantitative data analysis was done on SPSS version 27. I first performed a series of 3 x 3 x 3 mixed model Analyses of Variance (ANOVAs) with Grouping Method (no grouping-NG, random grouping-RG, similarity-based grouping-SG), Stream Type (travel, grocery, knowledge), and Trial (Trial 1, Trial 2, Trial 3) as fixed factors. I found Stream Type did not have any significant main effects, so I removed this variable in the following analyses.

Interviews after the final survey were audio-recorded and then transcribed. Following an inductive and interpretive approach, I initially created codes from the quotes, then identified high-level themes and links between themes. I went through an iterative process of generating and merging the themes to refine the codes.

Results

In this section, I first examine whether Grouping Method affects Ease of Following the Chat (H1). Then I present the results of Conversational Enjoyment (H2), Liking of the Live Stream

(H3), and Interpersonal Attraction (H4). Following that, I answer research questions about participation (R1), mental workload (R2), and mental engagement (R3).

Ease of Following the Chat

H1 hypothesized that it is easier to follow the chat when watching live streams within a subgroup than without a subgroup. I ran a mixed model ANOVA with Grouping Method and Trial as independent variables and Ease of Following the Chat as the dependent variable. I found a main effect of Grouping Method ($F [2, 73.77] = 135.97$, $p<.001$). Participants reported that chat was more difficult to follow in the NG condition ($M = 2.71$, $SE = .17$) than in the RG ($M = 5.14$, $SE = .17$, $p<.001$) or SG ($M = 5.40$, $SE = .17$, $p<.001$) conditions. I did not find a significant difference between the SG and RG conditions ($p = .16$). I also found a main effect of Trial ($F [2, 66.00] = 6.26$, $p < .001$). Chats in Trial 3 ($M = 4.81$, $SE = .18$) were easier to follow than chats in Trial 1 ($M = 4.21$, $SE = .18$, $p = .01$) and Trial 2 ($M = 4.22$, $SE = .18$, $p = .001$). I did not find a significant interaction between Group Method and Trial ($F [4,92.86] = 1.74$, $p = .15$). H1 is supported. The results are shown in Figure 7.

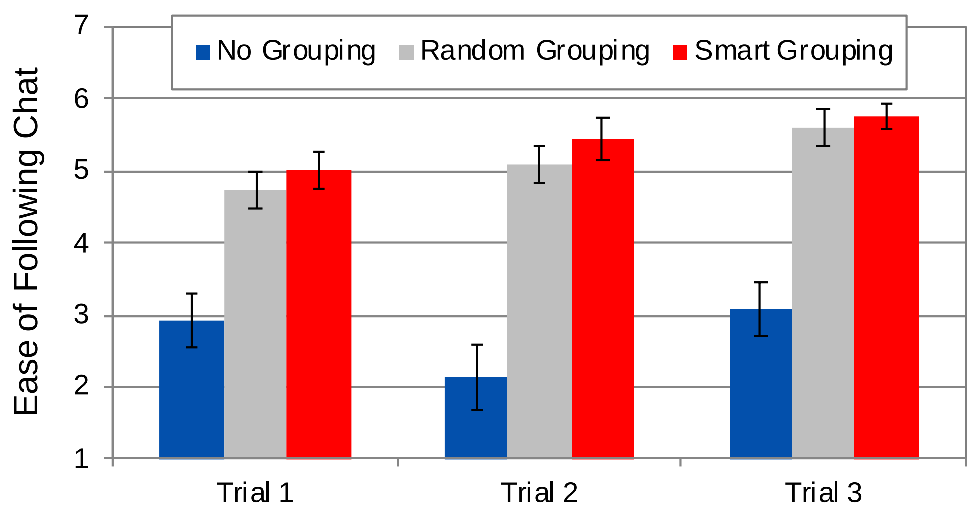


Figure 7: Ease of Following the Chat across three trials and grouping methods on a scale of 1 (low) to 7 (high). Error bars represent the standard error of the mean

Conversational Enjoyment

H2 hypothesized that viewers' conversational enjoyment will be positively associated with the ease of following chat conversations. I first conducted a mixed model ANOVA with Grouping Method and Trial as fixed factors and Conversational Enjoyment as the dependent variable. I found a main effect of Grouping Method (F [2, 67.72] = 6.01, p = .004). Participants reported that conversations in the RG (M = 3.55, SE = .21, p = .04) and SG (M = 3.84, SE = .21, p = .001) conditions were more enjoyable than conversations in the NG condition (M = 3.07, SE = .21). I did not find a difference between the SG and RG conditions (p = .20). Results also indicated a main effect of Trial (F [2, 60.10] = 3.85, p = .03), such that conversations in Trial 3 (M = 3.83, SE = .21) were more enjoyable than those in Trial 1 (M = 3.11, SE = .21, p = .01). There were no significant interactions (F [4, 93.46] = .86, p = .49, see Figure 8).

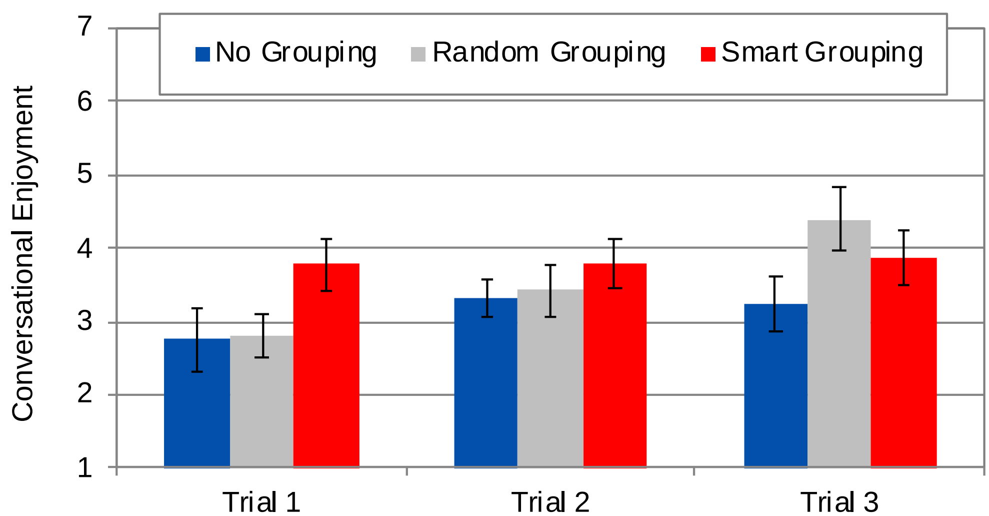


Figure 8: Conversational Enjoyment across three trials and grouping methods on a scale of 1 (low) to 7 (high). Error bars represent the standard errors of the mean.

I then added Ease of Following the Chat as a covariate to the mixed model as I hypothesize it would mediate the effects of Grouping Method. I found that Ease of Following had a significant effect on Conversational Enjoyment (F [1, 94.86] = 7.28, p = .01). However, with Ease of

Following in the model, the effects of Grouping Method (F [2, 82.48] = .65, p = .52) and Trial (F [2, 62.25] = 2.37, p = .1) were no longer significant. This indicates that Ease of Following the chat mediates the effects of Grouping Method and Trial on Conversational Enjoyment. Hence H2 is supported.

Participation

RQ1 asked whether ease of following the chat would affect viewers' participation in live streams. I ran a logistic regression with Grouping Method, Trial, and Ease of Following the Chat as independent measures and the binary participation variable as the dependent measure. I did not find any effects of Grouping Methods (B = .48, p = .15), Trial (B = -.01, p = .96), or Ease of Following the Chat (B = -.15, p = .38). Thus, participants' chatting behaviors were not affected by Grouping Method or Ease of Following the Chat.

Liking of the Live Stream

H3 hypothesized that viewers' liking of the live video would be positively associated with conversational enjoyment. I ran a mixed model ANOVA with Grouping Method and Trial as fixed factors and Liking of the Live Stream as a dependent variable. There was no main effect of Grouping Method (F [2, 69.12] = .19, p = .83) but a significant effect of Trial (F [2, 65.48] = 9.05, p < .001), such that the liking of the third live stream was higher (M = 4.90, SE = .26) than the first (M = 3.44, SE = .26, p < .001) and the second one (M = 4.24, SE = .26, p = .03). I did not find a significant interaction between Grouping Method and Trial (F [4, 98.57] = .92, p = .45).

A further analysis added Conversational Enjoyment as a covariate into the model. I found Conversational Enjoyment had a main effect on Liking of the Live Stream (F [1, 90.05] = 34.85, p < .001). There was a main effect of Trial (F [2, 66.45] = 5.19, p = .01) but not Grouping Method (F [2, 72.29] = .53, p = .59) and no interaction (F [4, 97.00] = 1.71, p = .15). This indicates that

Conversational Enjoyment affects Liking of the Live Stream. H3 is supported. But the effects do not come from Grouping Method. There may be other factors influencing Liking of the Live Stream.

Interpersonal Attraction

H5 hypothesized that viewers' perceptions of other audience members would be positively associated with conversational enjoyment. I first conducted a mixed model ANOVA of Grouping Method and Trial on Interpersonal Attraction. Results showed that Grouping Method had a main effect (F [2,70.89] = 7.85, p = .001), where participants in the SG condition perceived other audience members as more attractive (M = 3.20, SE=.19) than those in the RG (M = 2.58, SE = .19, p = .002) and NG conditions (M = 2.50, SE = .19, p = .001). There is no significant difference between the NG and RG conditions (p = .69). I also found a main effect of Trial (F [2,62.19] = 7.59, p = .001), such that participants reported that other audience members were more attractive in the third live stream (M = 3.24, SE = .19) than in the first (M = 2.39, SE = .19, $p < .001$) and second streams (M = 2.66, SE = .19, p = .002). There was no significant interaction between Grouping Method and Trial (F [4, 90.69] = 1.56, p = .19, see Figure 9).

Next, I added Conversational Enjoyment as a covariate to the mixed model. Results showed a main effect of Conversational Enjoyment (F [1,97.99] = 93.51, $p < .001$), Grouping Method (F [2, 71.54] = 4.63, p = .01), and Trial (F [2, 62.41] = 4.19, p = .02). The interaction between Grouping Method and Trial was also significant (F [4, 92.68] = 2.45, p = .05). This result indicates that Conversational Enjoyment mediates the effects of Grouping Method and Trial on Interpersonal Attraction, supporting H4.

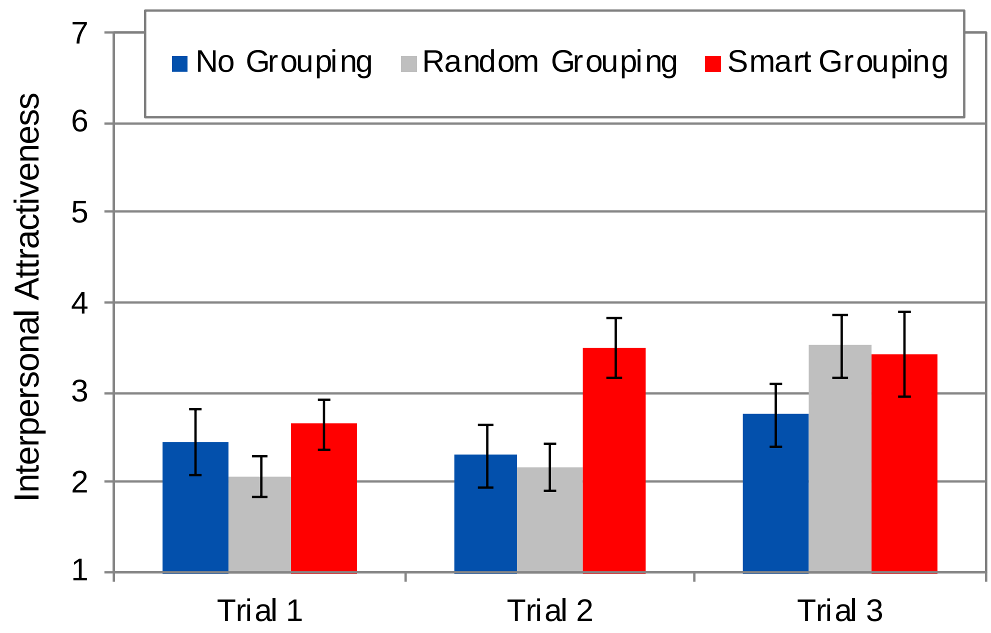


Figure 9: Interpersonal Attraction on a scale of 1 (low) to 7 (high) across three trials and grouping methods. Error bars represent the standard errors of the mean.

Mental Workload and Mental Engagement

RQ2 asked how subgrouping based on similarity affected participants' mental workload when interacting with live streams. To answer this question, I ran a 3 x 3 x 3 mixed model ANOVA with Grouping Method, Trial, and Stream Type as independent variables and Mental Workload as the dependent variable. I did not find a main effect of Trial (F [2, 45.95] = .13, p = .88) but there were borderline significant effects of Grouping Method (F [2, 53.18] = 2.53), p = .09) and Stream Type (F [2, 53.18] = 2.84, p = .07). Participants reported that mental workload was higher in the NG condition (M = 2.71, SE = .18) than in the RG condition (M = 2.31, SE = .18, p = .03). But the SG condition (M = 2.49, SE = .18) was not different from the NG (p = .22) or RG (p = .32) conditions. Participants also reported that the knowledge live stream (M = 2.71, SE = .18) required more mental workload than the travel stream (M = 2.28, SE = .18) (p = .02).

RQ3 investigated how subgrouping based on similarity affected participants' mental engagement when interacting with live streams. I performed a similar mixed model ANOVA using Grouping Method, Trial, and Stream Type. The results showed that Grouping Method did not have a main effect ($F [2, 51.09] = .56$, $p = .57$). But there were significant results of Trial ($F [2, 48.33] = 5.18$, $p = .01$) and Stream Type ($F [2, 51.09] = 6.80$, $p = .002$). Participants reported that the third live stream ($M = 4.41$, $SE = .22$) was more engaging than the first ($M = 3.46$, $SE = .22$, $p = .002$) and the second ones ($M = 3.81$, $SE = .22$, $p = .03$). The Knowledge live stream ($M = 4.47$, $SE = .22$) was more cognitively demanding than the travel ($M = 3.55$, $SE = .22$, $p = .001$) or food live streams ($M = 3.65$, $SE = .22$, $p = .004$).

The Influence of Attention on the Chat

During the interviews, I found some participants reported they did not pay much attention to the chat. Because this study was conducted online, I was not able to quantify the effects of visual attention to live stream chats using technologies such as eye-tracking. I asked the participants to share the screen and reminded them to pay attention to the live video and chats.

Although only a few participants sent messages to the chatroom, most of them reported reading the chat when watching the streams. Some of them reported reading less or ignoring the chat because of the large chat volume or because they were attending to the video instead. It is hard to separate the effects of chat from the entire live stream watching experience because the presentation and content of the chat may affect the willingness to read the chat. A participant might first read the chat and find the content overwhelming, confusing, or distracting, then decide not to read any more of the chat.

To investigate whether visual attention to the chats affects the results, I removed eight cases based on participants' self-report in the interviews. The results were basically the same as I reported above. There is a main effect of Grouping Method on Ease of Following the chat ($F [2, 67.13] = 107.81$, $p < .001$), Conversational Enjoyment ($F [2, 62.91] = 3.58$, $p = .03$), and Interpersonal Attraction ($F [2, 64.67] = 5.68$, $p = .005$). The main effect of Grouping Method on Conversational Enjoyment disappeared after including Ease of Following the Chat as a mediator ($F [2, 76.23] = .91$, $p = .41$) while the effect of Ease of Following the Chat is significant ($F [1, 87.59] = 7.27$, $p = .008$). The main effect of Grouping Method on Interpersonal Attraction still exists after including Conversational Enjoyment as a mediator ($F [2, 64.85] = 3.75$, $p = .03$). There is no main effect of Grouping Method on Mental Workload ($F [2, 67.24] = 1.71$, $p = .19$) and Mental Engagement ($F [2, 64.71] = .40$, $p = .68$).

Discussion

I proposed a grouping method that created subgroups based on live stream viewers' similarities. I examined the effect of the grouping strategy on participants' behaviors and perceptions in a laboratory study. Table 5 summarizes the hypotheses and results. Viewers reported enjoying the conversations more within a subgroup, mediated by Ease of Following the Chat. Viewers saw other audience members as more attractive when they are in a similarity-based subgroup, mediated by Conversational Enjoyment. I did not find a main effect of Grouping Method on Liking of the Live Stream.

In this section, I first discuss the study results and connect with previous literature, then I envision the design of future AI-embedded CMC with regard to subgroup formations and ethical issues. I talk about limitations and future work in the end.

Hypothesis	Supported
H1: Following conversations with a subgroup is easier than without a subgroup when watching live streams.	Yes
H2: Viewers' conversational enjoyment will be positively associated with the ease of following conversations.	Yes
H3: Viewers' liking of the live stream will be positively associated with conversational enjoyment.	No
H4: Viewers' perception of other members will be positively associated with conversational enjoyment.	Yes

Table 5: Hypothesis and Summary of Results

Participants reported chats in SG and RG conditions were easier to follow and more enjoyable. This result is consistent with previous studies (M. K. Miller et al., 2017). Yet I did not find a significant difference between the RG and SG grouping methods. It appears that the effects on Conversational Enjoyment may come from the slower chat speed rather than similarity of interests. However, as shown in the interview findings, some participants commented that chats in the SG condition were more meaningful and more relevant to the live video than those in the random grouping condition.

I did not find a significant effect of Grouping Method on participants' liking of the live stream. According to the model, Conversational Enjoyment contributes to the degree participants like the live stream. However, it is hard to claim there is a causal effect. During the interviews, when asked about the general impression of the live stream and chats, some participants mentioned that they paid more attention to the chats because they found the video content interesting and wanted to know how other people felt about it.

RQ2 and RQ3 asked about the effects of subgrouping on viewers' mental workload and engagement. I found the SG condition was not significantly different from the NG or RG conditions. Since the survey asked about the experience of watching the live streams as a whole, it is hard to separate the effects of video vs. chat. The results do suggest that subgrouping did not distract viewers or add to their mental workload. Future work can use physiological measurements such as biosensors and eye-tracking to capture workload and engagement more accurately.

The interviews reveal that participants often shifted their attention between the live video and chats. Sometimes they found the chats distracting, especially when they want to focus on the video. Other times they found the chats helped them understand the video and bring more fun to the watching experience. Watching live streams and chatting with others at the same time may involve a type of “channel blending” in which people intuitively reallocate mental resources between the video channel and text chat channel to best enjoy the watching experience (Isaacs et al., 2012).

Results showed that participants’ chatting behavior did not depend on Grouping Method or Trial (RQ1). This result is not surprising. People’s online behaviors are affected by a variety of factors, such as personality traits (Craker & March, 2016), social networks, or platform design (McInnis et al., 2016). As some participants mentioned in the interview, they generally prefer to read content generated by others instead of engaging with other users online. Some participants did want to join in the chat but could not think of anything valuable to contribute. In addition, as a short-period study that introduced participants to a new virtual environment, it is natural that participants did not find a personal connection and preferred to observe. Therefore, participation may not be a good measurement of positive live stream experiences.

I found that the effect of SG on interpersonal attraction was significantly stronger than RG. Participants perceived other audience members as more attractive when they were in the SG vs. the RG condition. I found that the effect of SG on interpersonal attraction was significantly stronger than RG. Participants perceived other audience members as more attractive when they

were in the SG than in the RG condition. This finding is novel in social psychology as previous studies either used perceived similarity (Maloney-Krichmar & Preece, 2005) or let members interact face-to-face (Harrison et al., 2002). My study created similarity-based subgroups and let participants construct an impression of others based on anonymous chat. This indicates that this grouping method not only offered a solution to overwhelming chat but may be applied to support community building. Previous literature on live streams emphasized the concept of community as it creates a virtual third space to socialize and participate (Hamilton et al., 2014), motivates the streamer, and keeps the stream sustainable (J. Li et al., 2019). In a recent study, Sheng and Kairam (2020) found that assumed common ground, and elevated visibility is important to form new relationships on Twitch; the relationships go stronger when there is frequent communication and move to intimate communication modes. This finding aligns with social psychology literature that the process from weak tie to strong tie needs repeated exposure, social interaction, and self-disclosure (Kraut & Resnick, 2012). Staying in a subgroup where people share similar attributes may help members to initiate social interactions, strengthen relationships, and develop personal connections. The increased bond-based attachment may contribute to community maintenance.

The study design has several limitations. First, I primed the participants on which part of the video they should pay attention to as their “interest”. In practice, people usually watch the live streams they are interested in. This manipulation may reduce ecological validity, which is a common problem for lab experiments. As I did not find a main effect of stream type, future work can select one live stream and recruit real followers as participants. Second, I simulated the live stream watching experience instead of letting participants engage with real live streams. Although this method controls variables such as chat content, it may fail to capture the full interaction experience. In addition, the time frame of the study may be too short for participants to get involved in the task. In the future, I plan to explore the effects of smart subgrouping in a real-world scenario. Third, the subgrouping methods in this study focus on viewers' experience but does not consider what the streamer will see. Subgroups based on similarity may be able to generate a theme or topic which is easier for a streamer to interpret. Future work should explore how to design for streamers if subgrouping is adopted in live streams.

CHAPTER 5

CHATBUDDIES: AN INTELLIGENT USER INTERFACE EMBEDDED WITH SUBGROUPING ALGORITHMS IN LARGE-SCALE LIVE STREAMING

Findings from the controlled lab experiment in Chapter 4 suggest that subgrouping is an effective approach to lower the difficulty of following the chat and increase conversational enjoyment. Furthermore, similarity-based grouping is positively associated with interpersonal attraction such that participants were more attached to other viewers. These findings indicated the potential of similarity-attraction theory in large-scale real-time communication. In this chapter, I present Chatbuddies, a framework and an intelligent interface that generate subgroups based on viewer similarity.

Design Iterations

Chatbuddies was designed using Figma, a design-based vector graphics and prototyping tool. I followed an iterative design process such that each version of the design was informed by user feedback. Seventeen participants were recruited from an institution participant pool and received course credits as compensation. Each participant was presented with a few interactive design prototypes in random order and was invited to interact with the prototypes for a few minutes. They were asked how they think of the design, the aspects they like or dislike, and their preferences among the prototypes. Figure 10 shows some design prototypes and feedback.

Iteration One (in Figure 10) splits chats into two windows: global chat and local chat. Global chat shows all the chat messages as a normal live stream. Local chat displays chat messages that are relevant to the user based on similarity. Users can view the explanation of global chat and local chat by clicking the “question” icon. They can see a full list of viewers by clicking the “people” icon in the top right corner. Chat messages were retrieved from a real live stream and the usernames were pseudonymous. Participants commented that Local Chat provides customized content, but it was confusing to see replicated messages in both windows. Splitting chats left and right makes it difficult to read chats. Furthermore, participants reported that they did not understand how the subgrouping algorithm changes the chat groups.

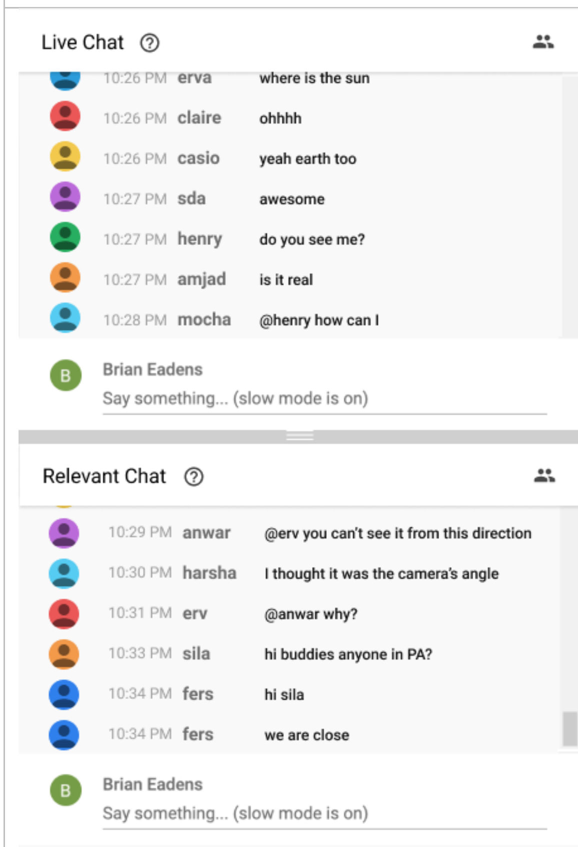
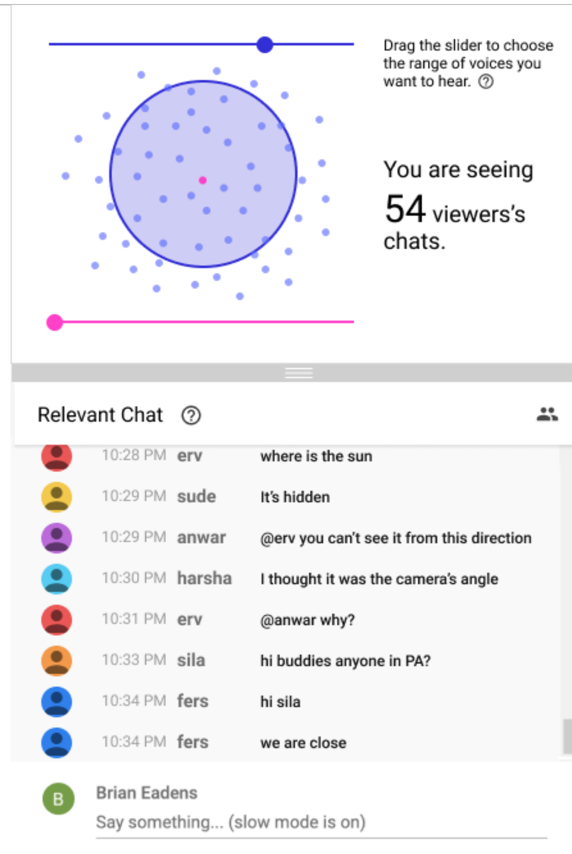
<table>
<tr><th>Iteration 1A: Split up and down</th><th>Iteration 2: Visualization panel</th></tr>
<tr><td>
</td><td>
</td></tr>
<tr><td>Dislikes: (1) Not sure what "local" means; (2) hard to follow chats in both windows, especially in Iteration 1B; (3) confused when seeing replicated messages in both windows; (4) do not have a sense how the algorithm changes chat groups</td><td>Dislikes: (1) Confused about the functionalities of the two sliders and two circles at the first glance. (2) not sure the directions of sliders (larger or smaller;</td></tr>
</table>

Figure 10. UI of Iteration One and Two. Iteration One: Global chat shows all the chat messages as a normal live stream. Local chat displays chat messages that are relevant to the user based on similarity. Iteration Two: A visualization panel with a scatter plot shows the current user and other viewers.

Iteration Two (in Figure 10) added a visualization panel, in which a scatter plot represents all the viewers. The besides text explains that users can drag the sliders and control the volume of messages. Participants reported that it was a bit hard to understand the functionalities of the two sliders at the first glance, which was figured out after clicking the "help" icon.

Symmetric and Asymmetric Grouping

Two types of subgrouping strategies were generated in the design process: Symmetric and Asymmetric Grouping (Figure 11). This is an essential design question in group communication. The concept of symmetricity is based on the information viewers obtain in the chat. In symmetric grouping, each viewer is in one single subgroup, and they can see each other's messages, like a traditional group chat. In Asymmetric Grouping, however, each viewer has their own "neighborhood" (M. K. Miller et al., 2017) and they can see messages sent by peer viewers who are the most similar to them. As a result, they may not see the entire conversation of others, the information is not symmetric among all the viewers.

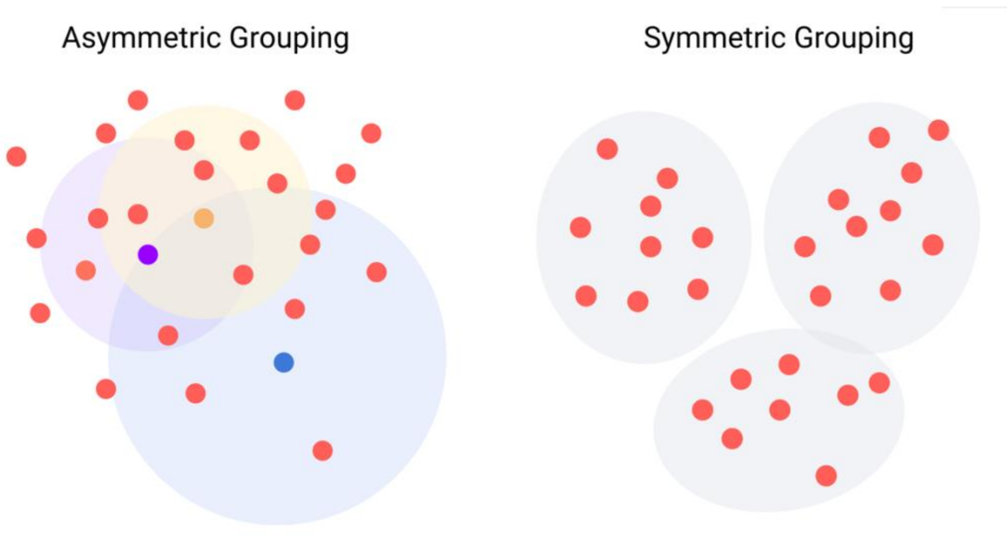


Figure 11. The Graph Concept of Asymmetric and Symmetric Grouping

The two strategies are generated based on two user needs: have a meaningful conversation with a smaller group of viewers, view messages that are relevant to them. Both design strategies have been applied in CMC technologies, such as Facebook Group (symmetric), Zoom breakout rooms (symmetric), News Feed (asymmetric), etc.

I experimented with more iterations on the two types of subgrouping strategies and collected more feedback. Iteration Three to Five are demonstrated in Appendix C.

Iteration 6A: Asymmetric grouping	**Iteration 6A: Adjusted circle size**	**Iteration 6A: Explanation after clicking "learn more"**

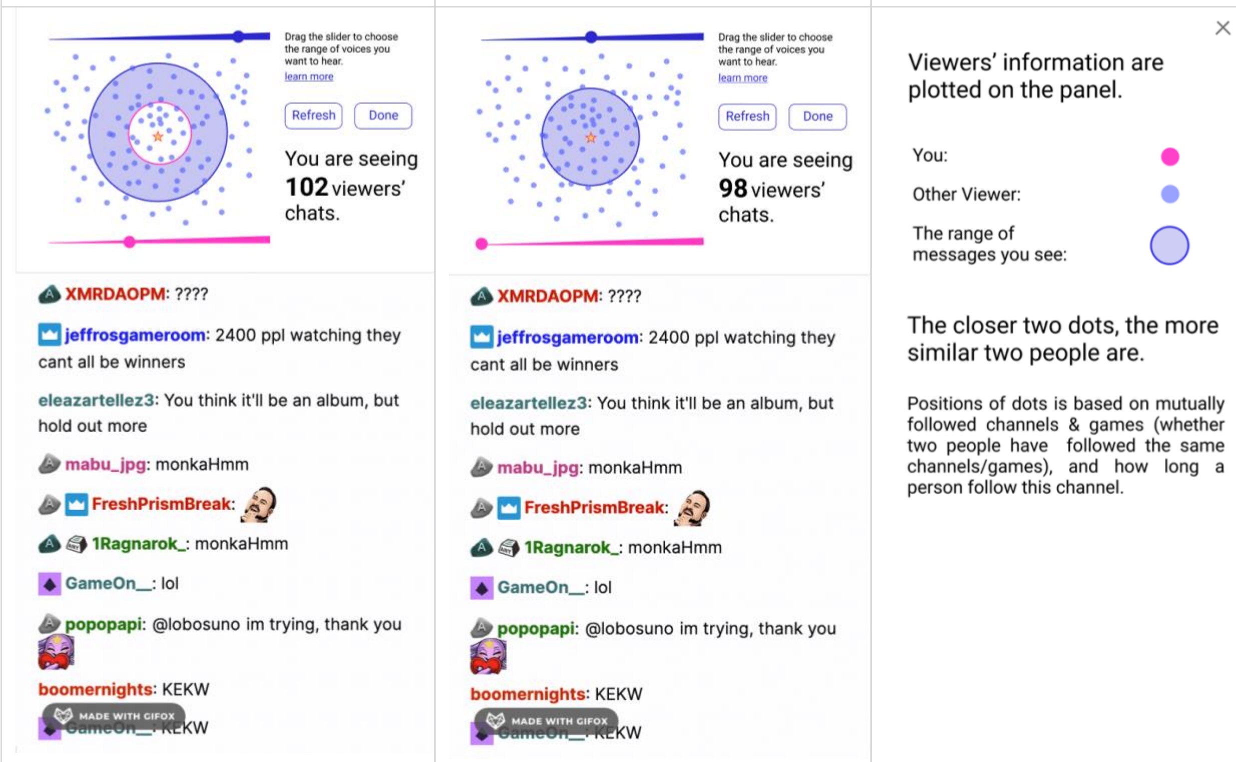


Figure 12. Iteration 6A: Asymmetric Subgrouping with User Control.

Iteration Six was demonstrated in Figure 12 and Figure 13. On Asymmetric Grouping interface (Figure 12), a "subgroup" is created around an individual user. The star at the center of the graph represents the current user. Users can adjust the volume of chat messages by dragging the sliders. Sliders will be hidden if it is not in edit mode. Users can see an explanation board that demonstrates the meaning of the scatter plot and features used in the subgrouping algorithm.

On Symmetric grouping interface (Figure 13), each user is assigned to one single group and groups are color-coded. The current user is highlighted with a star icon.

In general, I found most people like the idea of displaying chat messages sent by people whom they may be interested in, such as those in the same time zone or those who have similar interests. They preferred a clear explanation of why a certain group of chatters is relevant to them and a representation of the grouping structure of the chatroom when subgroups are generated.

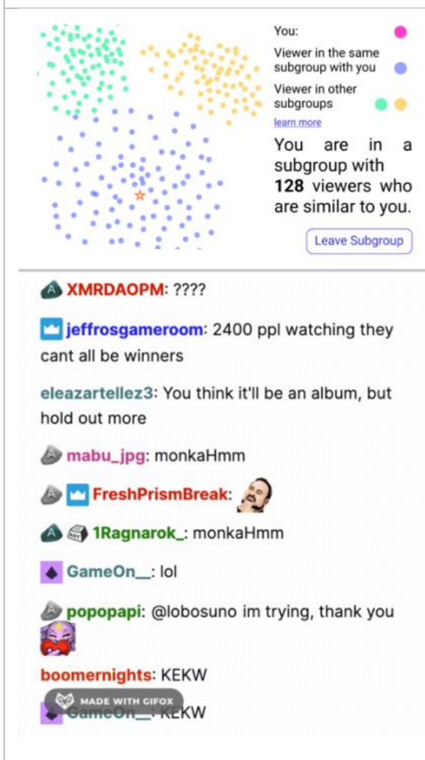
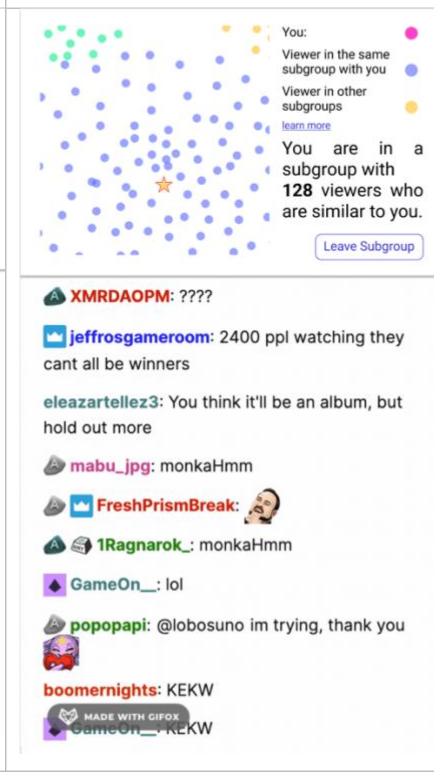
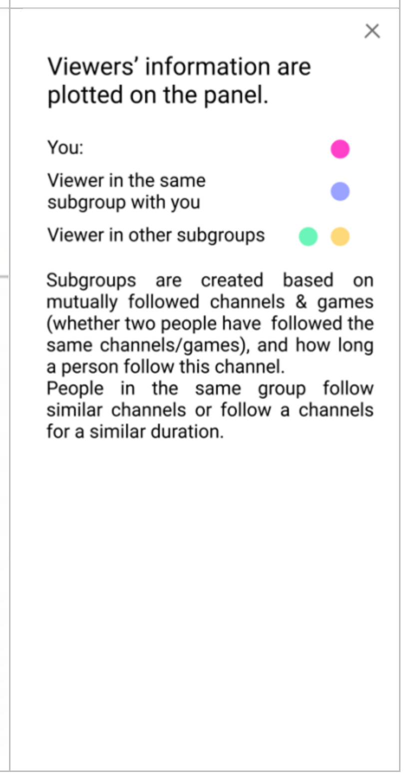
Iteration 6B: Symmetric grouping	**Iteration 6B: visualization panel zoom in**	**Iteration 6B: Explanation after clicking “learn more”**
		

Figure 13. Iteration 6B: Symmetric Grouping

Proposed Framework

Step 1: Detect Subgroup Need. Not all the live streams need subgroups. In small streams where a smooth conversation is not a problem, or in streams where most viewers prefer to see all the chats, there is no need to introduce subgroups. An arbitrary method is to monitor the number of concurrent viewers or the chat speed. Once the audience size comes across a threshold, or the chats reach a speed that makes it difficult to read chats, the next step will be triggered.

Step 2: Generate Subgroups. Chatbuddies backend server first queries viewers' information, such as the subscription status of the current channel and following channels (details in the next section), then returns the grouping results to the client side.

Step 3: Visualize Grouping Results. Once the client side receives outputs from the remote server, it displays the subgrouping results on a scatter plot and demonstrates the meaning of the graph. Viewers see chats sent by people in the same subgroup.

Chatbuddies – An Intelligent Live Stream Chat Interface

I have demonstrated the design process of subgrouping interfaces. In this section, I describe Chatbuddies, an intelligent interface that displays relevant chat messages in large-scale live streams. Chatbuddies is designed as a framework that embeds subgrouping algorithms and data visualization into a live stream chat interface, instead of as an alternative to current live stream systems. It can be implemented as an extension or add-on to integrate into any live stream platforms, such as Twitch and YouTube. I present the framework of the system, the algorithms embedded and how algorithms are selected, the interfaces and features.

System Overview

Figure 14 illustrates the pipeline of Chatbuddies. The frontend chatroom sends a subgroup request to the server, then the server retrieves a viewer list and viewer data from live stream API

(Twitch API is used in this dissertation). The backend server uses viewer data to generate subgroups and sends results back to the frontend. The chatroom message list and visualization panel render the graph and display chat messages according to the group assignment.

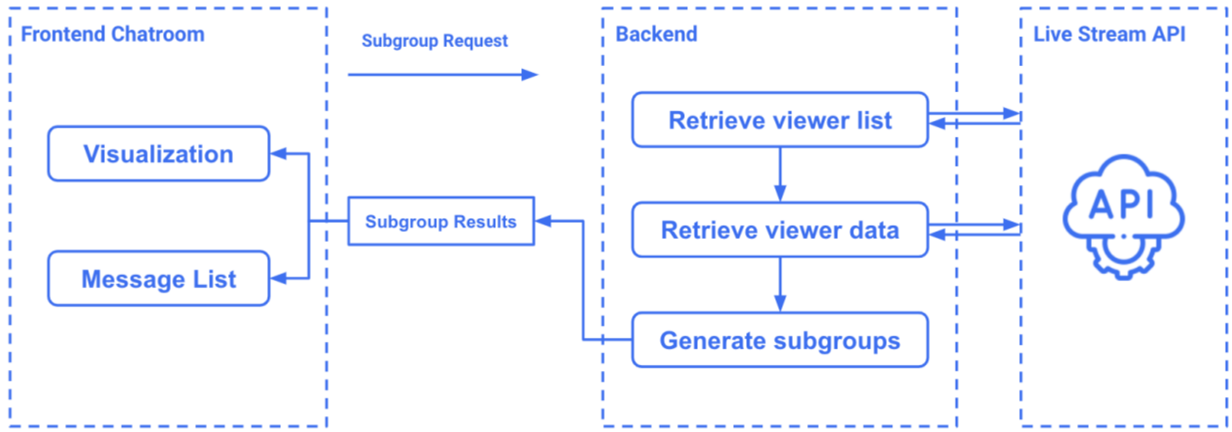


Figure 14. Chatbuddies System Pipeline.

Subgroup Generation Algorithms

In Chapter 4, subgroups were manually generated by selecting a subset of chatters whose chat messages largely regard a particular target topic, for instance, if more than 60% of a chatter's messages are about the scenery in a travel live stream, the chatter was assigned to the scenery topic subgroup. In practice, however, systematically generating subgroups for live stream chatrooms is far more nuanced.

In this section, I go through feature selection, data preprocessing, clustering algorithm selection, and evaluation. Clustering algorithms are applied only in symmetric mode. Algorithms used in asymmetric mode will be demonstrated later.

Feature Selection

Features (similarity dimensions) are a set of measurable properties that represent a particular user such that each user is an n-dimensional vector in the feature space. Features are selected based on two principles: efficacy of information and viability of implementation. As a result, features used in the algorithms should be supported by either literature or empirical evidence. Features should also be reliably retrievable. Previous research studying the similarity-attraction phenomenon investigated a variety of attributes such as personality (Banikiotes & Neimeyer, 1981; Byrne et al., 1967; Duck & Craig, 1978), attitudes (Byrne, 1961), hobbies (Curry & Emerson, 1970; Werner & Parmelee, 1979), demographics (Baxter & West, 2003; McCroskey et al., 2006; Sprecher, 1998) and more. Certain attributes may not be applicable in live streaming contexts, such as demographic information that is missing on anonymous platforms or protected attributes that are ethically dubious to retrieve.

According to Byrne's reinforcement theory and information processing theory, some attributes provide more information about an individual and are more important features (Montoya & Horton, 2012). Members of a video game community may not care about others' political leaning or opinions on organic foods.

Therefore, I used two kinds of features: common ground and channel seniority. Common ground can be understood as mutual knowledge between individuals, such as commonly followed channels and commonly watched games. Sheng & Kairam (2020) found that common knowledge contributes to the formation of friendships on Twitch. Channel seniority indicates how long an individual watches a certain channel. A senior follower may have followed a channel for many years and is very familiar with the streamer. A newbie may just find a channel and follow for several days.

I retrieved data from Twitch API: 1) a list of channels a user follows; 2) games or categories of the stream most recently broadcasted by these channels; 3) how many days a user follows a channel; 4) a user's subscription tier of a channel (1000, 2000, 3000 or not subscribe).

Extracting Viewer Characteristics.

Given the first two features are in the format of lists, I apply two approaches to further process the data, Binary and Affinity Graph. The first approach is to convert the channels and games a viewer follows to a binary value. For example, 525 viewers who are watching a live stream at the same time follow 1324 channels in total. For each channel, if a viewer follows it, I mark it as 1, otherwise as 0. The same data processing is applied to the following games feature. The number of features is determined by the number of channels and games viewers follow.

# of features = # of channels viewers follow in total + # of games viewers follow in total + 1 (duration of following) + 1 (subscription tier)

Another approach is to compute a similarity score between two viewers. Given two lists of channels two viewers follow, I measure the overlap between the two viewers via Jaccard similarity (Tan PN, Steinbach M and Kumar V). The same measurement is applied to games viewers follow.

The similarity of following duration and subscription tier is measured as follows. I compute the differences of days two viewers follow a channel and map the difference over a certain period, e.g., ten years. For example, Viewer A has followed Channel X for 1320 days, Viewer B just follows Channel X for 45 days. The similarity score will be *1 – days_diff / max_diff = 1 - (1320-45) / (10 * 365) = 0.65*. If two the differences are larger than ten years, the similarity score is close to 0. If a viewer does not follow a channel, the value will be *-1 * max_diff* therefore this viewer will be dissimilar to those who follow Channel X.

Twitch has three subscription tiers (Tier 1, Tier 2, and Tier 3) with different costs. The similarity score of this feature is mapped to three tiers. For example, Viewer A subscribes Channel X at Tier 1, Viewer B subscribes at Tier 3, their *similarity score = 1- (3-1) / 3 = 0.33*. If one of the viewers does not subscribe but the other does, the similarity score is 0. If both two viewers do not subscribe, the score is 1.

The computation above gives us four similarity scores corresponding to the four features retrieved from Twitch API. The scores are integrated with weights to compute the final similarity score between two viewers. Weights options are evaluated in the later sections. In the end, we can get an $N * N$ affinity matrix where N = the number of viewers in the current channel.

Clustering Algorithms

There are many clustering algorithms, each having its own strength and weakness, based on the data characteristics. In the proposed system, I consider four commonly used clustering algorithms, partitioning, density-based, hierarchical, and graph-based algorithms (Rai & Singh, 2010; D. Xu & Tian, 2015). I did not include deep learning network algorithms because of their black-box nature and their high computational requirements. As a proof concept, I anticipate classical clustering models should be robust enough.

I selected KMeans, KMedoids, Density-Based Spatial Clustering of Applications with Noise (DBSCAN), Agglomerative, and Spectral Clustering as candidate algorithms. KMeans is one of the most popular clustering algorithms and has been applied in lots of clustering scenarios (Deshpande et al., 2020). But its performance suffers from high-dimensional and binary data because its similarity measurements, Euclidean distance, "reduces to counting the number of variables on which two cases disagree" (IBM, 2020). "After the initial centers are chosen (which depends on the order of the cases), the centers are still binary data. For the first iteration, as the cases are compared to cluster means, they will always be at some integer distance from each of the centers. There will often be ties, and the case will be assigned to a cluster in an arbitrary manner"

(IBM, 2020). I perform a Principal Component Analysis (PCA) on the Binary data structure first and apply KMeans.

KMedoids is another partitioning algorithm but different from KMeans in the definition of center and measure of similarity. Kmedoids uses actual data points as centers instead of an average. It can use arbitrary dissimilarity measures instead of Euclidean distance in KMeans. In the evaluation, a precomputed affinity matrix is used as input data.

DBSCAN (Ester et al., 1996) is designed to discover clusters of arbitrary shape by finding the high-density areas in the domain and expanding those areas of the feature space around them as clusters. It does not require a pre-defined number of clusters and can identify outliers as noises. But it performs poorly when the clusters are of varying density.

Gaussian Mixture Models (GMMs) assume data points are Gaussian distributed, which is a more flexible assumption compared with KMeans that data points are circular. It uses probability to assign clusters, therefore one data point can have multiple clusters. But it requires lots of parameters and many iterations to get good results (Mohammed et al., 2016).

Arising from the concepts in spectral graph theory, the basic idea of spectral clustering is "construct a weighted graph from the initial data set where each node represents a pattern and each weighted edge simply takes into account the similarity between two patterns" (Filippone et al., 2008). I use the Affinity Graph data structure as a fully connected graph with different weights.

Algorithm Evaluation

To evaluate the effectiveness of these clustering algorithms in Chatbuddies, I conduct an experiment applying the clustering algorithms above.

I first retrieved members from 15 discord servers. Then I collected user information based on the features described above, namely, channels followed, most recently games played by the following channels, whether they follow a certain channel, and how long they follow a channel.

I first paired two discord servers, resulting in 105 unique pairs. For each pair, I randomly selected 100 members from each server and mixed them as a list of 200 members. Then I applied the algorithms to generate two clusters. I used Normalize Mutual Information (NMI) as the measurement as it is widely used in clustering task evaluations if the "ground truth" is available (Schütze et al., 2008). NMI measures the overlapping between the generated clusters and ground truth.

Table 6 presents the algorithm evaluation results. Four weights options were evaluated for the affinity matrix as input data where the value of weights indicated the importance of three features: channels a viewer follows, games played by following channels, and the duration of following a channel. Feature subscription status was removed because this data is only available with user permission. Some cases with a zero mean were removed. In general, Spectral Clustering reached the highest NMI score, though the value is relatively low. Weights with [0.2, 0.2, 0.6] were selected as it has a balanced mean and standard deviation.

Algorithm	Feature Type	Weights	Mean	SD
Random Clustering	NA	NA	0.00	0.01
Dimensional Reduction plus Gaussian Mixture	Binary	NA	0.01	0.01
Dimensional Reduction plus KMeans	Binary	NA	0.01	0.01
KMedoids	Affinity Matrix	[0.2, 0.2, 0.6]	0.02	0.02
KMedoids	Affinity Matrix	[0.33, 0.33, 0.33]	0.02	0.02
KMedoids	Affinity Matrix	[0.1, 0.1, 0.8]	0.03	0.04
Agglomerative Clustering	Affinity Matrix	[0.4, 0.4, 0.2]	0.05	0.05
Agglomerative Clustering	Affinity Matrix	[0.33, 0.33, 0.33]	0.06	0.06
Agglomerative Clustering	Affinity Matrix	[0.2, 0.2, 0.6]	0.08	0.07
Spectral Clustering	Affinity Matrix	[0.4, 0.4, 0.2]	0.11	0.09
Agglomerative Clustering	Affinity Matrix	[0.1, 0.1, 0.8]	0.11	0.09
DBSCAN	Affinity Matrix	[0.2, 0.2, 0.6]	**0.15**	0.36
DBSCAN	Affinity Matrix	[0.33, 0.33, 0.33]	**0.15**	0.36
DBSCAN	Affinity Matrix	[0.4, 0.4, 0.2]	**0.15**	0.36
Spectral Clustering	Affinity Matrix	[0.33, 0.33, 0.33]	**0.18**	0.11
Spectral Clustering	Affinity Matrix	[0.2, 0.2, 0.6]	**0.29**	0.17
Spectral Clustering	Affinity Matrix	[0.1, 0.1, 0.8]	**0.32**	0.24

Note: some cases with a zero mean were removed.

Table 6: Mean NMI of clustering algorithms. Algorithms with the highest mean NMI are Special Clustering, DBSCAN, and Agglomerative Clustering.

Chatbuddies Interface

Building on the feedback I got in the iterative design process, I developed, Chatbuddies, a live stream chat interface that generates subgroups for viewers. The frontend web interface is built using React.js and D3.js. The backend server is built with Flask and Firestore database. The web app can be launched in a variety of browsers such as Chrome, Firefox, Safari, and Edge.

Visualization of Subgroups

Chatbuddies is a framework that embeds subgrouping algorithms in the live stream chat, therefore it should be compatible with most live stream platforms. Figure 15 is the experiment website I used in the evaluation experiment. Besides basic chat functions, Chatbuddies has a viewer visualization component at the top of the chat channel. This design decision comes from the formative interviews that people want to understand the communication organization if subgroups exist in a huge channel. Furthermore, such visualization can provide an interactive approach for users to explore the grouping decision. Some intelligent systems present algorithmic results as a black box instead of offering explanations. But studies showed that interactive and white-box explanations improve users' comprehension of the system (H.-F. Cheng et al., 2019). I follow a transparent AI principle (Abdul et al., 2018) such that systems should inform and explain to users its algorithms.

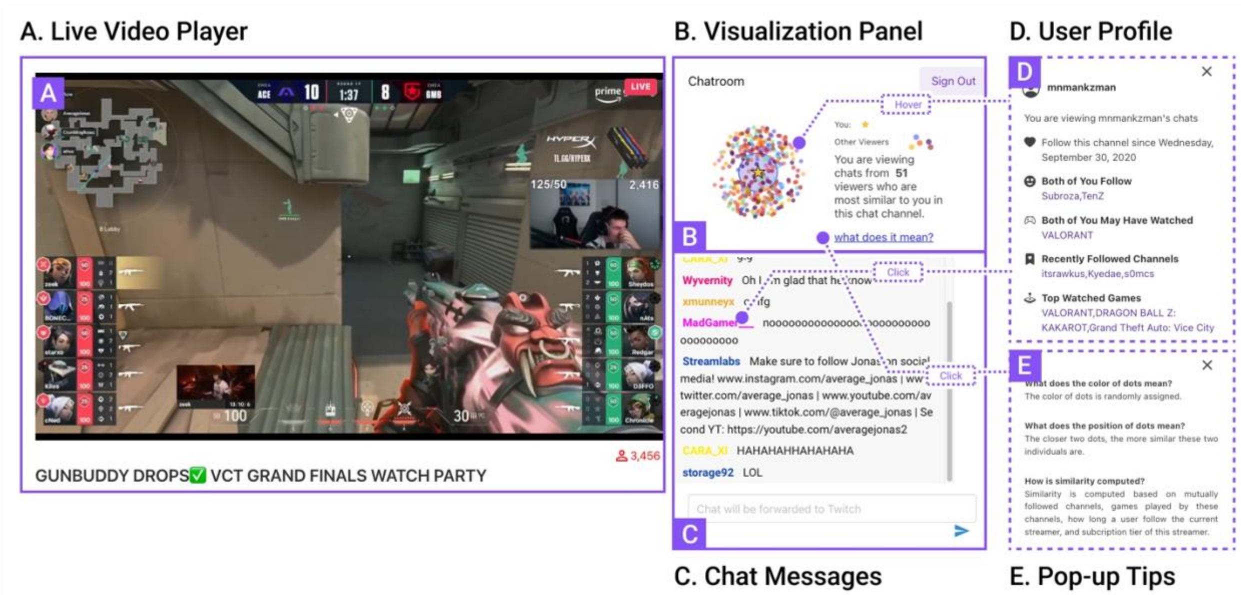


Figure 15: Chatbuddies Interface (Asymmetric Similar).

(A) Live video player. (B) Visualization panel where individual viewers were presented as scatter plots. (C) Chat messages. (D) A user profile is triggered by either hovering on the dot or clicking the username in the chat message. (E) Pop-up tips that explain the grouping algorithm.

In the visualization panel, viewers in the current stream are illustrated in a scatter plot. The "Visualization Panel" column in Figure 16 demonstrates the interfaces in symmetric and asymmetric modes.

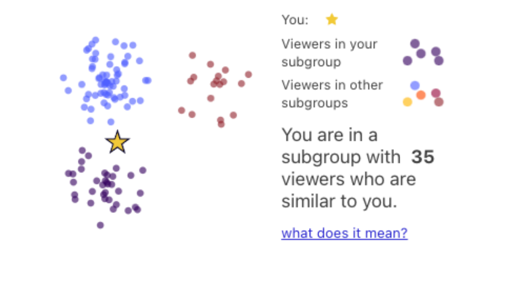
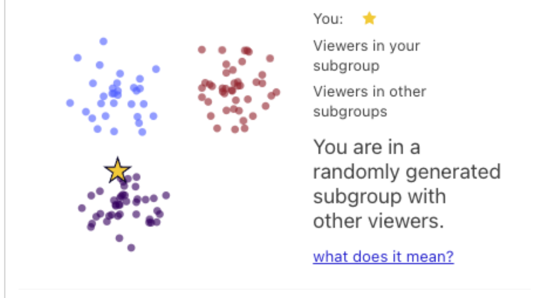
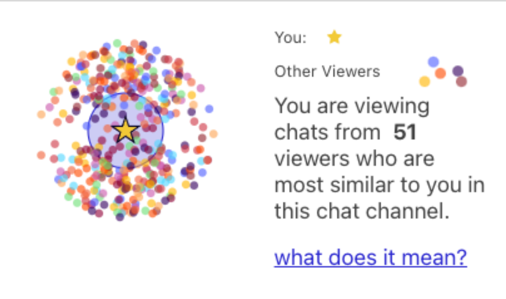
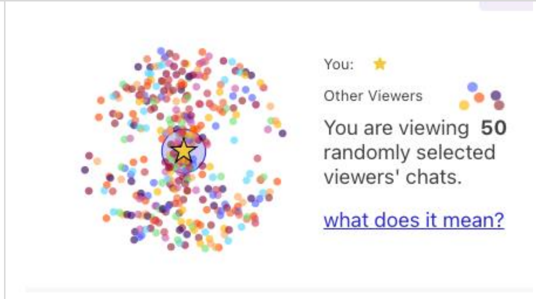
	Similar	Random
Symmetric		
Asymmetric		

Figure 16: Visualization panel in four conditions.

In symmetric clustering mode, a viewer can only see messages sent by people in the same subgroup. The scatter plot shows viewers in groups. Each group has a unique color. Both the cluster center and the position of individual dots are randomly initiated. Remember Chatbuddies applies Spectral Clustering where the affinity instead of Euclidean distance between each viewer pair is manually computed. Therefore, it is not easy to plot all the dots with their distances representing affinity on a 2D graph.

In asymmetric mode, each user has a unique subgroup. Users can view messages sent by a certain number of people who are the most similar to them in the current live stream channel. As described in the last section, similarity is computed based on four features describe in the section above: channels they follow, games played by these channels, the duration they follow the current channel, and their subscription status. The number of viewers whose messages are displayed to a user is determined based on the stream size and chat speed. By default, the value is between 40 and 50 based on previous studies (M. K. Miller et al., 2017).

Users can learn more about the graph and algorithms by clicking "What does it mean?" on the visualization panel, and a pop-up tip will show up and explain the algorithm, as shown in Figure 15.

Visualization of Individual Viewers

When a user hovers over the dot or clicks the username in the chat messages, the viewer's information shows up. The information includes the viewer's username, the duration they follow this streamer, channels both the viewer and the user follow, games played by the commonly followed channels, channels recently followed by the viewer, and the games played by the channels.

Onboarding Tutorial

User interviews in the prototyping phase indicated that new users found it hard to fully understand the system and graph at the first glance. Therefore, new users are instructed by a step-by-step onboarding tutorial about the meaning and functionalities of each interface element. Onboarding contents are demonstrated in Figure 17.

More details of Chatbuddies interfaces can be found in Appendix D.

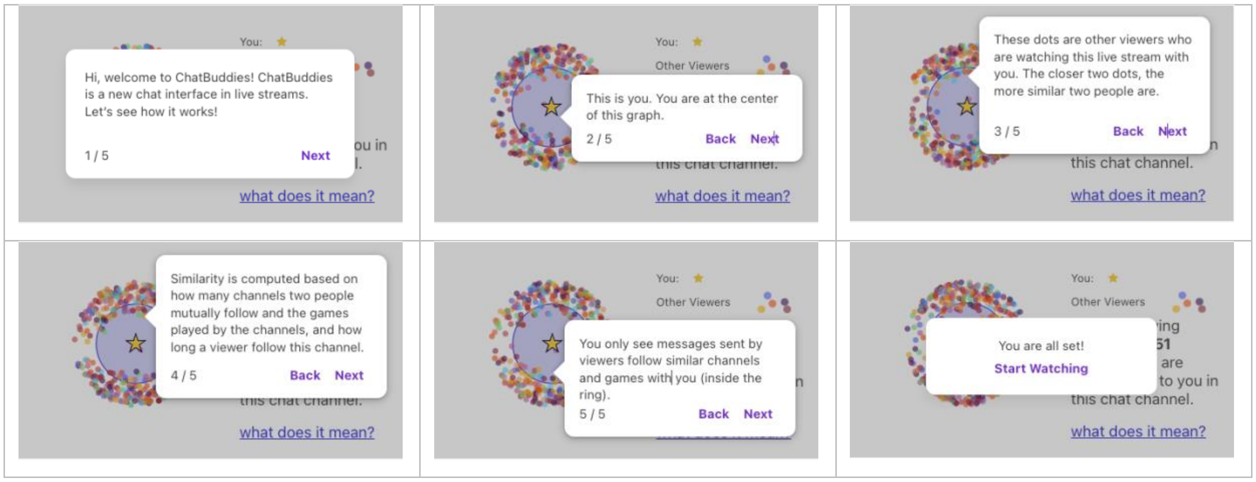


Figure 17. Onboarding tutorial steps from left to right, top to bottom

Evaluation

To evaluate the effects of the proposed subgrouping framework, I designed an online unmoderated experiment. The goal is to 1) understand how active live stream users interact with this interface and 2) compare the effects of similarity-based grouping and random-based grouping.

In this section, I first discuss the justifications of evaluation approaches. Next, I present instruments, participants, and the study procedure. Then I elaborate on measurements and data analysis approaches.

Evaluation Method Rationales

The evaluation has the following challenges: 1) Participant recruitment. Chatbuddies is a framework that generates subgroups based on user data. In this project, I used Twitch as an applied platform. To evaluate such a framework, participants need to be Twitch users, which is a niche population. 2) Experiment setup. The most ecologically valid evaluation approach is to apply this framework in a real large-scale live stream, where many participants watch a stream and then give feedback as in Miller et al. (2017)'s study. It requires lots of participants to join the study at the same. Given the target participants are already a niche demographic, it is not likely hundreds of people who are willing to participate in a study are available at the same time.

Therefore, a simulated live stream approach was applied. Participants watched a prerecorded live stream and read chats sent when the stream was live. They were told that the system randomly picked a Twitch live stream for them to watch, they can view others' chats from Twitch and their chats will be forwarded to Twitch. In actuality, they were not chatting with other viewers as the stream was not really live. This approach has been used in previous studies (Maruyama et al., 2017). Though it introduces some bias if the participant recognizes the video was not live and they were not chatting with real viewers, it provides a stable experiment environment as the video content and chats are controlled. To simulate the interactions and control the impact of reciprocity, I used the same technique in Chapter 4 that participants received at most two messages with “@<username>” once they participated in the chat. Chat content was selected from a prepared pool of messages that shows greetings, such as “hi”, and “good morning.”

Applying the simulated “live” stream technique to Symmetric Grouping evaluation is riskier than Asymmetric Grouping because viewers in one group should be continuous and perform as having a conversation. Displaying a subset of chat messages retrieved from a past live stream does not meet this criterion; as a result, the deception technique may not be successful. Whereas Asymmetric Grouping has the nature of personalizing chats, so the loss of frequent interaction is not expected. The study only evaluates the Asymmetric Grouping technique with two conditions:

- **Asymmetric Similar**: create asymmetric subgroups based on similarity
- **Asymmetric Random**: create asymmetric subgroups randomly

Research Questions

For the evaluation study, I ask the following questions:

RQ1: How do participants behave on Chatbuddies Asymmetric Similar and on Asymmetric Random interface?

RQ2: How do Chatbuddies Asymmetric Similar and Asymmetric Random interface affect participants’ Conversational Enjoyment and Interpersonal Attraction?

RQ3: How do participants perceive and compare Chatbuddies subgrouping algorithms?

Participants

I recruited participants from the following sources: Amazon Mechanical Turk, social media ads, Facebook groups, and authors’ personal social networks. Qualified participants meet the following requirements: 1) over 18 years old; 2) located in the US; 3) regularly watch gaming live streams; 4) own a valid Twitch account. I want to recruit participants that are best fit the target users who are likely to watch a gaming live stream.

A total of 72 participants were recruited and received compensation at a rate of $16 per hour. Eight participants were excluded from the analysis because either they requested to remove their data after reading the debriefing statement at the end of the study, or because the data quality was low (e.g., short response time, answers to all the questions are the same). This left 64 participants aged from 20 to 60 (Mean = 32.66, SD = 8.13), including 18 women, 45 men, and one reported as non-binary.

I report their demographics for the sake of completion. Among the 55 self-reported as non-Hispanic, 46 are White, 13 are Black, three are Asian; two reported as other. Around half (N=30) of the participants reported they watch live streams every day; 20 participants watch more than once a week. Most participants have participated in live chat (N=62); 27 reported as frequent chatters (selected “always”, “most of the time”). In terms of the size of the stream, 5 participants reported that they usually watch small live streams where concurrent viewers are fewer than 100 viewers, 29 watch streams with around 100 to 1,000 viewers; 23 watch streams with 1,000 to 10,000 viewers; seven watch streams with more than 10,000 viewers.

Materials

Chatbuddies Live Stream Chat System

To simulate the context of how a Twitch user normally watches a live stream, an unmoderated online experiment approach was employed. Participants can join the study any time they want, and their behaviors were not observed by the experimenter.

The following features were added to the system to adjust the experiment needs and make it an online unmoderated experiment.

Authentication. Participants log in to the system by typing their Twitch username. The web application immediately validates the account through Twitch API by checking if the account exists, if the account has registered for more than a month, and if the account follows at least three channels (otherwise there is too little information to generate groups). Originally Chatbuddies asks for Twitch API authentication by redirecting to Twitch and logging in. But it takes time and effort, and some users have concerns with third-party authentication, which may be a barrier to recruiting participants.

Embedded Surveys. Two Qualtrics surveys were embedded into the experiment website. The first one is a consent form with some general questions about participants' live stream watching experience. Another one is the final survey which automatically shows up once the participants have watched the "live stream" for 15 minutes. The details of the surveys will be described below.

User Behavior Monitor. In an unmoderated web-based experiment, it is hard to guarantee participants stay in front of the computer and watch the entire stream. To maximize the experiment validity, the following strategies were used to monitor participants' behaviors. First, before watching the "live" stream, participants were reminded that they ought to stay on the site before the time is up. If they do not have enough time today, they can come back any time they want. Second, the experiment website embeds a recorded Twitch stream and chats retrieved from Twitch. It also stores the watching progress of each participant. If a participant leaves the current webpage, the video will pause. It is to avoid the case that participants just open the tab and do other things on the browser, which is a common use of live stream for background noise (Taber et al., 2020).

When participants come back, the webpage will remind them how long they have left, and the video will resume with several seconds after where it pauses. It has a risk that participants recognize it is not "live" if they leave for hours. But in reality, very few people did so.

Live Stream and Chats

I requested and was granted permission from a streamer to use two recorded live streams as the study material. The streams were recaps of VALORANT[3] competitions, where the streamer and his friends watched the competition together and commented about their reactions and personal experience. The original live stream lasted for more than five hours. Two 15-minute clips were selected to use in this study. The clips were selected because they were close to the end of the competition and include highlight moments, for example, the players have a good shot or something unexpected happened so that participants would not feel bored in the study. The content, atmosphere, and excitement of the two clips were similar based on the feedback of a pilot study (N=14). I retrieved the chats from the original stream through Twitch API. I went through the chats and removed chats with inappropriate words. The usernames of these chats' authors were modified when displayed to participants.

Surveys

Each participant completed two post-session surveys and one post-study survey. Participants were directed to complete a post-session survey after they have watched the "live" stream for 15 minutes. The post-study survey showed up after the participant finished the second post-study survey. Post-session surveys include questions about the watching experience, such as liking of the live stream, attention division, conversational enjoyment, etc. Post-study survey asked

[3] VALORANT is a 5v5 character-based tactical first-person shooting game. https://playvalorant.com/en-us/

participants' general watching habits, (i.e., how often they watch gaming live streams on Twitch, how often they send chat messages when watching gaming live streams), the motivation of watching gaming live streams, the functionalities of chats in their mind, the stream size of the live streams they usually watch. A full list of questions can be found in Appendix E.

Procedure

Participants first logged in via their Twitch username. Then they were directed to a consent form that explained the study purpose, procedure, requirements, compensation, and data storage. Participants who granted consent to participate in the study will continue to answer a few questions about their general watching experience, including frequency of watching gaming live streams on Twitch, frequency of sending chats in Twitch gaming live streams, the size of live streams they usually watch, the motivation of watching Twitch gaming live streams, the function of chats in their mind. Then participants were told the system was picking a Twitch live stream for them to watch. They were told they can freely express their reactions (M. K. Miller et al., 2017). They were told that they have to continuously watch for 15 minutes to count full participation. If they don't have time at this moment, they can do the study when they have time. After 15 minutes, the experiment website showed a survey with questions about their liking of the live streams, attention split, event awareness, conversational enjoyment, interpersonal attraction, and perceived similarity. Participants were also asked a manipulation check question about what subgrouping algorithm was used in the study and two open-ended questions about their thoughts on this live stream system and ideas for improvement.

Participants were told to participate in the second session one day later. The second session was scheduled one day later instead of immediately after the first session because I used live stream videos from the same streamer. It is not common that a streamer changes stream content in a short time. But it is common that one streamer broadcasts every day on the same topic or video game. Participants recruited through MTurk were invited to join the second session through the MTurk system. Participants recruited via social media were invited through email. The procedure was the same as the first session except one more open-ended question asking them to compare the two subgrouping algorithms were asked at the end of the second session.

Measurements

Two types of measurements were used in this study: self-report survey and behavioral logs.

Survey Measurements

These questions are asked in the post-session survey.

Attention Split. I want to understand how the grouping principle and criteria affect users' behaviors when watching live streams. Participant were asked how they split the attention between video and chat. Options are "fully on video", "mostly on video", "half video half chat", "mostly on chat", "fully on chat."

Event Awareness. To future understand the attention split and engagement, I asked the participants to indicate the extent to which they were aware of the events in the video and topics in the chats. Questions are "I was aware of most events in the live stream" and "I was aware of most messages in the chat." Participants indicated on a 7-point scale ranging from "Strongly Disagree" to "Strongly Agree".

Conversational Enjoyment. Four items based on Ryan & Deci (2000) about interest and enjoyment were adapted. Questions included: "I had so much fun participating in this conversation when watching the live stream", "I thought the conversation was enjoyable", "I enjoyed the conversation with others when watching this live stream", and "The conversation did NOT hold my attention at all (reserved)". Participants indicated on a 7-point scale ranging from "Strongly Disagree" to "Strongly Agree". Factor Analysis with Varimax Rotation indicated that all items fell on a single dimension accounting for 71.58 % of the variance, so I averaged them to construct this measurement (Cronbach's alpha =.94).

Interpersonal Attraction. Items were adapted from similarity effects literature and online communities (Sprecher, 2014; Ren et al., 2002), including 2 items about affective attraction ("I felt close with them when watching this live stream", "I would like to be friends with them") and two items about behavioral attraction ("I am interested in learning more about them", "I like these people"). Participants indicated on a 7-point scale ranging from "Strongly Disagree" to "Strongly Agree". Factor Analysis with Varimax Rotation indicated that all the items were on a single dimension. Therefore I averaged them to create the measure of interpersonal attraction (Cronbach's alpha =.93).

Perceived Similarity. Specific similarity in gaming preference, chat style and personality were asked in the format of "How much do you think you have in common with the other chatters in terms of the following dimensions?" Participant answered on a 7-pointed scale ranging from "Not at All" to "a Great Deal". Factor Analysis with Varimax Rotation indicated that the three items were on a single dimension. The three items were averaged to create a general perceived similarity (Cronbach's alpha =.93). Note that this question was asked after all the other questions to avoid bias.

Behavioral Measurements

Number of Profile Card Clicks. The number of times a user clicks either their own or other viewers' profile card to learn more information. The clicks include those triggered on the scatter plot and those triggered from message list.

Number of Explanation Card Clicks. The number of times a user clicks "what does it mean?" and read the explanation.

Number of Chats Sent. The number of messages a user sent during each live stream.

Onboarding Tutorial Reading Duration (seconds). The number of seconds a user spent from the first time they interact with the onboarding text to finish the tutorial.

Data Analysis

All statistical analysis were performed using SPSS software version 27. Since it is a within-subject experiment, I conducted 2 x 2 Mixed Model ANOVA and took Grouping Method condition (Asymmetric Similar, Asymmetric Random) and Trial (Trial 1, Trial 2) as fixed factors, and multiple hypothesized factors as covariate. The term "trial" and "session" are used interchangeably, meaning the 15-min live stream plus a post-session survey.

Open-ended questions were analyzed to better understand potential reasons of their response and their suggestions for future design. There are two open-ended questions in the post-session survey asking for comments of the interface and ideas for improvements. Another open-ended question was asked after the participants finished two sessions about their perceptions of the two interfaces they have used. I first open-coded the responses and then identified similar responses and grouped them together to generate themes. This process iterated for several rounds to merge themes and redefine codes.

Results

In this section, I first present users' interactions with Chatbuddies interface. Next, I report the results of manipulation check and the hypothesis. Then findings of qualitative data are discussed.

Manipulation Check

Besides a simple manipulation check question "How did the system select the subset of viewers", perceived similarity in gaming preference was also used as a measurement of manipulation check. A 2 x 2 Mixed Model ANOVA showed that the perceived similarity of gaming preference in Asymmetric Similar condition (M = 4.66, SE = .23) was higher than Asymmetric Random condition (M = 4.17, SE = .23, $p < .01$). Trial does not have a main effect ($p = .20$).

Interaction with Chatbuddies Interface

Behavioral measurements are summarized in Table 7. Most of these measures have a skewed distribution.

Measurements		Mean	SD	Min	Max	Median
Onboarding Tutorial Reading Duration		19.02	14.04	3.80	75.97	15.61
Number of Explanation Card Clicks		.18	.44	0	2	0
Number of Profile Card Views	(Total)	10.83	19.39	1	96	3
Number of Profile Card Views	(Self, Panel)	1.05	3.06	0	19	0
Number of Profile Card Views	(Self, Chat)	.02	.12	0	1	0
Number of Profile Card Views	(Other, Panel)	10.25	18.41	1	92	3
Number of Profile Card Views	(Other, Chat)	.11	.42	0	3	0
Number of Chats Sent		4.80	5.27	0	35	4

Table 7. Behavioral Measurements Results

In 32.0 % (41/128) of sessions participants read tutorials within 10 seconds, in 50.78% (65/128) of sessions they read tutorials between 10 to 30 seconds. Participants spent more time on reading the tutorials in Asymmetric Similar condition (M = 21.81, SE = 1.69) than in Asymmetric Random condition (M = 16.12, SE = 1.72; $F[1, 62.86] = 6.89$, $p = .01$). They also spent more time in the first session (M = 21.62, SE = 1.70) than the second session (M = 16.31, SE = 1.70; F [1, 61.60] = 6.03, p = .02).

In most sessions (84.40%) participants did not click "What does it mean?" for further information. It can be understood as either the onboarding tutorials explained the features well or participants were not interested in learning more about the subgroup generation. Due to the data sparsity of this measurement, it was excluded from the later analysis.

Participants did spend time learning about the peer viewers by viewing their profile card. On average in each 15-min session participants viewed others' profile three times. I converted the Number of Profile Card Views (Other, Panel) to a binary variable such that participants who viewed others' profiles on visualization panel more than median were classified as viewed (1), otherwise as not viewed (0). A mixed model ANOVA found a borderline effect of Trial (F[1, 62] = 3.54, p = .06) but did not find a main effect of Grouping Method (F [1, 62] = .1.05, p = .31).

In most sessions (83.59%, 107/128) participants joined the chat conversation. In 44.5%, (57/128) of sessions participants sent more than five chat messages in the channel. There was no difference of number of chats between the two conditions (F [1, 62.18] = .24, p = .63). Because this variable is not normally distributed, it was recomputed as a binary data such that cases with at least one chat sent were classified as "participated", and those without chats were classified as "not participated." The percentage of participation was 84.38% (54/64) in Asymmetric Similar condition and 82.81% (53/64) in Asymmetric Random condition.

Attention and Awareness

Besides system log data, participants' self-reported behavior data were also analyzed. A Mixed Model ANOVA showed that there was no significant difference in attention split between the two conditions (F [1, 62.88] = .01, p = .91). Grouping Method has a borderline significant effect of on awareness of events ($F [1, 62.73] = 3.76, p = .06$) and a main effect on awareness of chats ($F [1, 63.01] = 6.62, p = .01$). Participants reported to pay more attention on both the video and chats in Asymmetric Similar condition (video: $M = 5.66, SE = .17$; chat: $M = 5.45, SE = .17$) than in Asymmetric Random condition (video: $M = 5.32, SE = .18$; chat: $M = 4.92, SE = .18$). Trial has no main effect on awareness of video ($F [1, 62.09] = .52, p = .47$) but has a borderline effect on awareness of chat ($F [1, 62.07] = 2.95, p = .09$).

Conversational Enjoyment

I conducted a mixed model ANOVA with Grouping Method and Trial as fixed factors and Conversational Enjoyment as the dependent variable. Main effects of Grouping Method ($F [1, 62] = 14.48, p < .001$) and Trial ($F [1, 62] = 4.95, p = .03$) were found. Participants reported that they enjoyed the conversation more in the second trial ($M =4.48, SE = .21$) than the first trial ($M = 4.17, SE = .21$); and they enjoyed the conversation more in Asymmetric Similar condition ($M = 4.59, SE = .21$) than in Random condition ($M = 4.06, SE = .21$).

Interpersonal Attraction

A mixed model ANOVA was conducted with Grouping Method and Trial as fixed factors and Interpersonal Attraction as dependent variable. A main effect of Grouping Method was found ($F [1, 62.10] = 2.05, p < .001$), but no main effect of Trial ($F [1, 62.21] = .39, p = .53$). Participants reported that they feel more attracted by viewers when using Asymmetric Similar interface ($M = 4.60, SE = .21$) than Asymmetric Random interface ($M = 4.12, SE = .21$). Results are shown in Figure 18.

A body of literature argued the effect of perceived similarity on interpersonal liking (Sprecher, 2014; Sprecher et al., 2015; Sunnafrank, 1986). To examine it I added Perceived Similarity as a covariate to the model, Grouping Method ($F [1, 64.71] = 2.81, p = .10$), Trial (F [1, 60.73] = .23, p = .64), and their interaction (F [1, 60.73] = .30, p = .59) were not significant anymore. Perceived Similarity has a main effect (F [1, 110.93] = 237.58), p < .001). The results indicate that Perceived Similarity mediate the effect of Grouping Method on Interpersonal Attraction.

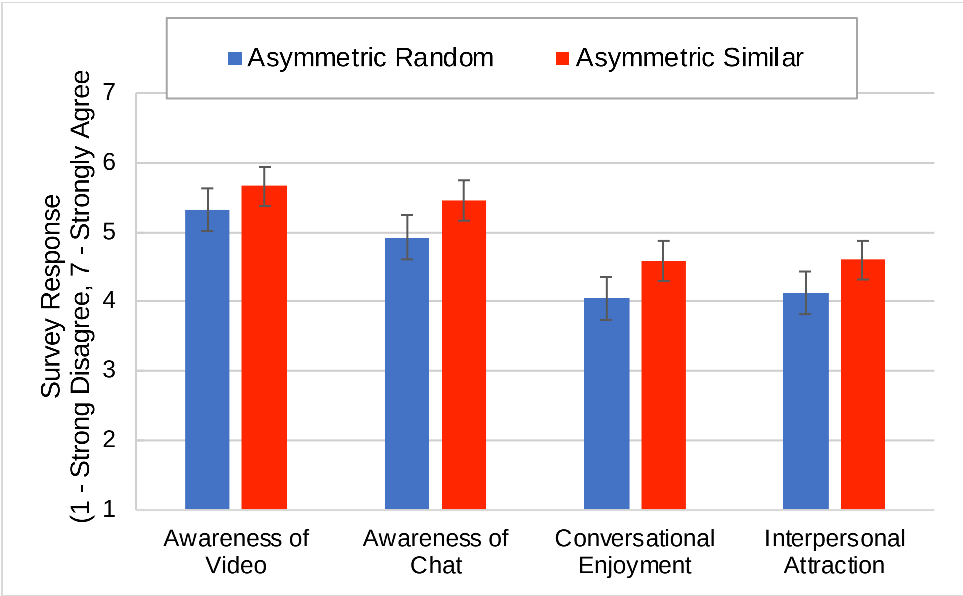


Figure 18. Means (±SE) for questionnaire questions on Grouping Methods on a scale of 1 (low) to 7 (high). Error bars represent the standard errors of the mean.

Qualitative Results

Responses to open-ended questions were analyzed using iterative open-coding and affinity diagramming (Holtzblatt & Beyer, 1997). Relevant responses were grouped together to generate a theme. The average word count of all the open-ended questions is 16.68. Most participants have positive comments on the entire system in terms of usability and ease of following the chat. Some of them commented on the underlining reasons for preferring one interface rather the other. I present the findings below. The quotes are presented in the format of (Participant number, Gender, Age).

Similarity-based grouping made it comfortable for people to chat. Participants reported that a smaller chatting group is more intimate, and subgrouping based on similarity makes it easier to find like-minded people.

> *"I liked knowing that the people in the chat were similar to me. It made me feel comfortable talking to them since I knew we enjoy the same things." (P8, Male, 39)*
>
> *"I think filtering the messages by similarity works better and is a better way to find like-minded people to chat with. I think the random method can work too, but it's easier to chat with people who are into the same games as you." (P9, Male, 34)*

Some participants mentioned talking to similar others is a "rewarding" experience.

> *"Similarity is much better and makes more sense to me. I'd rather be talking with people who I know that I have something in common with, rather than a random cross-section of the viewers as a whole. I think it would make for a more rewarding chatting experience." (P28, Male, 33)*

Some participants shared reasons for not preferring similarity-based grouping. The primary reason was the lack of diversity and chances to meet cool people. On the contrary, random grouping allows them to see messages from a diverse population and different angles.

> *"Personally I would prefer to see all of the messages in chat, not just from people with similar preferences as me. I think it's more satisfying to see a lot of messages from people with different backgrounds and interests." (P54, Male, 28)*
>
> *"It was better by random because you get to see different ideas and different comments that expound your thinking of the game and give you a more wider view of the game from different angles rather than similar which only gives you messages similar to yours." (P25, Male, 22)*

Some participants offered ideas for balancing the two algorithms and improvements. For example, the system can shuffle groups so that one user does not always stay in one group.

> *"Maybe instead of limiting to the 50 or so viewers most similar to me, find a way to rotate that group of 50 viewers. The way it is now, I might miss a chance to connect with another cool viewer since they might not meet that threshold of being similar enough to me." (P10, Male, 27)*
>
> *"Perhaps a feature where you can "shuffle" the chat the program places you in if one isn't liking those in their group." (P29, Female, 40)*

Another idea is to include more subgrouping dimensions other than those used in Chatbuddies, such as age and motivations for watching.

> *"It might be helpful if you could choose the type of people you'd be in a group with. For example, if you want to talk to people who are just goofing around, people who want to analyze the game, etc." (P50, Male, 48)*
>
> *"It would be nice to be able to sort or filter by your own metrics. Like age or something." (P31, Male, 37)*

Discussion

In this chapter, I present Chatbuddies, an intelligent live stream chat interface that creates subgroups to facilitate communication in large-scale live streams. The above results indicated that similarity-based asymmetric subgroup creations outperformed the random approach.

Participants reported that they were more aware of the chat, enjoyed the conversation more, and liked other chatters more when watching live streams on Asymmetric Similar interface than the Asymmetric Random one. In Chapter 4 I did not find a significant difference between similar-based and random grouping on Conversational Enjoyment. The inconsistency may be caused by the differences between experiment setup such as explicit information of grouping method and unmoderated and types of live streams. In the lab experiment, participants did not know how the subgroups were generated while in the Chatbuddies study participants were informed whether their subgroups were created by similarity or random.

Previous literature on similarity-attraction effects argued that it is perceived similarity instead of actual similarity that contributes to interpersonal attraction (Montoya et al., 2008; Sprecher, 2014; Sunnafrank, 1986). Instead of directly applying similarity-attraction hypothesis, I first carefully analyzed factors that may influence the link between the two variables (as shown in Chapter 2), I selected subgrouping algorithms and designed Chatbuddies interfaces based on the analysis. Key factors include similarity dimensions, manipulation approach, and interaction content. In this study, participants saw messages sent by similar others measured by common ground (commonly following channels and games played by these channels) and channel seniority (duration of following). Participants received information about asymmetric subgroup creation algorithms from an onboarding tutorial (required to read) and an explanation card (optional to read). Chatters in the live stream talked about the events in the live stream. Participants can send chat messages to the chatroom but were responded with blurry greeting messages only. It is worth noting that the goal of this study is not to examine which variable caused interpersonal liking, actual or perceived similarity. Results showed that both contributed to Conversational Enjoyment and Interpersonal Attraction and perceived similarity may mediate actual similarity. But the causal relationship was not clear since previous work reported enjoyment of interaction mediates the effect of perceived similarity on interpersonal liking (Sprecher et al., 2013). Without considering the underlying relationships among these factors, I applied the results of literature were such that both actual and perceived similarity was used: asymmetric subgroups are generated based on actual similarity; participants were informed that their peer viewers were similar to them or not. Participants can see channels and games followed by other viewers as the conversations in the live chat were going on.

There are several limitations of the evaluation study. First, as in above sections mentioned, the video and chats as well as the interaction were not really “live” therefore the study design can be understood as a modified version of bogus stranger technique. It is true that participants got little responses from other viewers and the information exchange was not as rich as in previous studies where dyads talked for 5 to 30 minutes (Sunnafrank, 1983, 1984, 1985; Sprecher, 2014). But notice that most participants did not really join the conversation even though the stream was perceived as live. Second, the study was unmoderated, there was no way to guarantee all the participants took the study carefully. Third, Chatbuddies only has access to a public API data, which may not be as robust as other data such as friend networks and subscription status.

CHAPTER 6

GENERAL DISCUSSION: DESIGN FOR AI-EMBEDDED CMC

This dissertation aims to understand challenges in large-scale interaction and explore design opportunities of combining algorithms and CMC tools to facilitate communication and build connections. My work dives deep into the live streaming context, where live video and text chat are integrated to support streamer to viewer and viewer to viewer interaction. A qualitative interview study was conducted to understand people's motivation, practices, and challenges when engaging with live streams. I proposed a similarity-based subgrouping method and conducted Wizard-of-Oz lab experiments to evaluate the effects on viewers' watching experience and interpersonal outcomes. Then I designed Chatbuddies, an intelligent live stream chat interface that generates subgroups and visualizes information about peer viewers. An online experiment was conducted to investigate the benefits and costs of such a design. In this chapter, I summarize the results of my previous studies and discuss the design implications of intelligent CMC.

Summary of Results

My dissertation follows a Human-Centered Design method (Boy, 2011; Cooley, 2000): I first construct a detailed understanding of the motivation, practices, and challenges when people engage with large-scale live streams; then I brainstormed design ideas to address the challenges and evaluated the prototype through a controlled lab experiment; based on the results, I further designed and implemented an intelligent live stream chat framework and conducted an online experiment.

Chapter 3 presents a qualitative interview study that explored the motivation of watching live streams on SNS, the practices of engaging with the streamer and peer viewers, and challenges people face when watching live streams on SNS. Interviews showed that people watch live streams with purposes along two sets of dimensions: social or individual, and entertaining or intellectual. They watch live streams on SNS because they want to enjoy interests, to get information, to stay updated, and to connect with others on popular social media platforms. Interviewees' challenges were closely relevant to their engagement style and communication approach. I identified two styles of live stream interaction: loyal followers who focus on and react to the streamer, and community players who actively engage with other viewers in the live stream. Participants tend to have a preferred engagement style but may change depending on the live stream category, content, their relationship with the streamer, and other factors. Loyal followers mainly complained about distraction and lack of direct interaction with the streamer, while community players mostly reported suffering from a lack of coherent and meaningful conversation and of a friendly communication environment.

Based on the findings in Chapter 3, Chapter 4 proposed a similarity-based subgrouping method that aims to divide viewers into subgroups based on how similar they are on certain dimensions, such as interests. A controlled within-subjects lab experiment using priming and Wizard-of-Oz techniques was conducted to compare the effects of three grouping methods, No Grouping, Random Grouping, and Similarity-based Grouping, on watching experience and interpersonal outcomes. Participants watched three simulated "live" streams and view chats in the chat channel when the stream was live. They were primed to pay attention to certain topics in the three "live" streams. Participants reported that the chats in subgroups were easier to follow and more enjoyable than chats in a channel without subgroups. They perceived peer viewers as more attractive in similarity-based subgroups than in random subgroups. No difference of mental workload and mental engagement among the three subgrouping methods was reported.

Chapter 5 proposed Chatbuddies, an intelligent live stream chat framework that generates subgroups based on similarity of common ground and channel seniority. Chatbuddies has two subgrouping strategies: symmetric and asymmetric mode. In symmetric mode, an individual viewer stays in a unique subgroup, and they can see chat messages from others in the same subgroup; in asymmetric mode, each viewer has their own "subgroup," and they can see chat messages sent by those who are most similar to them. Due to recruitment difficulties of evaluating the symmetric version (see Chapter 5), an online unmoderated within-subject experiment was conducted to evaluate the effects of asymmetric similarity-based grouping on watching experiences and interpersonal outcomes, using random grouping as control condition. Twitch users who regularly watch gaming live streams were invited to watch two "live" streams and chat with peer viewers using the Chatbuddies interface. Participants reported higher conversational enjoyment and interpersonal attraction towards peer viewers when using the similarity-based interface.

Design Considerations and Implications for Large-scale Communication

In Chapter 2, I reviewed previous literature on similarity-attraction effects, discussed its theories and critiques, and described the factors that may affect the evaluation results. Chapters 4 and 5 illustrate a theory-inspired design approach through which design elements are justified based on analysis of both theories and empirical research experiments. Results indicate that similarity-based subgrouping is a promising approach to improve interaction experience in large-scale communication in live streams.

In this section, I first discuss the advantages and disadvantages of subgrouping, next I go through key factors mentioned in Chapter 2 and their roles in design considerations of large-scale communication. Then I discuss potential ethical issues from three perspectives.

Advantages and Disadvantages of Subgrouping and Similarity-based Approach

As the results in Chapter 4 and Chapter 5 showed, subgrouping has the following advantages: 1) It makes chat messages easier to follow; 2) It makes the conversations more enjoyable; and 3) Similarity-based approach enhanced viewers' interpersonal attraction towards chatters in the same subgroup. The first two results indicate that subgrouping can be a solution to information overload and enhance experiences in large-scale communication, which aligns with previous studies (M. K. Miller et al., 2017). The third result showed us an opportunity of similarity-attraction effects that may be used in community construction and maintenance. As participants in Chapter 5 pointed out, they felt more comfortable chatting with like-minded people. Common ground, such as shared objects of reference, paves the way for establishing communication (Clark & Brennan, 1991). Subgrouping may have further benefits such as regulation, as identifiable group members tend to act in a more group-normative manner (Douglas & McGarty, 2001). Within a smaller-sized and consistent group, members might be more self-regulated and use more acceptable languages.

However, disadvantages also exist. Results of open-ended questions in Chapter 5 revealed that participants were aware of the downside of subgrouping and similarity-based strategies. 1) Creating subgroups, either by random or by any clustering algorithms, displays only a subset of chats, which restricts viewers to catch the overall conversations. For viewers who prefer a holistic view of live chats, subgrouping may jeopardize their viewing experience; 2) Specific subgrouping approaches have their own problems. The symmetric algorithm introduced in Chapter 5 faces the challenge of stabilizing subgroups as viewers leave and join a live stream channel and they may watch one live stream for a short time on average (Santora, 2022). The asymmetric algorithm suffers from losing conversation threads from chatters outside one's own subgroup. A more thorough design is needed to fix these issues, such as allowing viewers to trace threads. 3) Similarity-based subgrouping may cause ethical problems by creating echo chambers (Sunstein, 2018) and biasing viewers' opinions. I will discuss this topic in the upcoming sections.

Factors that Impact Design Decisions of Similarity-based Subgrouping

As discussed in Chapter 2, designers and engineers should think thoroughly on some key factors and decide whether, when and how to apply similarity-based grouping.

Whether and When to Apply Subgrouping. Not all the large-scale real-time interaction needs subgrouping. As demonstrated in Chapter 5, if a chat channel has many users but only a few are talking, there is no need for subgroups. But once the interaction suffers from information overload, some clustering strategies may be helpful. In some domains or communities, similarity-based grouping may not be a good design. For example, in a book sharing community whose goal is to encourage idea exchange and discussion, similarity-based subgrouping may cause replicated content and less contribution (Ludford et al., 2004).

Similarity Dimension. The selection of similarity dimension depends on the domain, community purpose, and type of member attachment. Chapter 4 used interests in subtopic as similarity dimension because it was accessible in archive live chats. Chapter 5 used common ground and channel seniority based on communication theories, previous similarity-attraction studies, and feasibility of data retrieval. Studies in Chapter 4 and Chapter 5 covered domains such as grocery shopping, travel, knowledge sharing, and gaming, where the interactions are more entertainment-oriented, and members have more affective attachment. However, in a need-based health community, when users come to seek information and get answers, besides creating subgroups based on categories of diseases, designers can also consider pair users based on experience and expertise.

Similarity Information Explanation. In Chapter 2, I discussed similarity manipulation approaches in previous literature. “Manipulation” means an experimenter pairs similar or dissimilar participants together based on some dimensions. In F2F communication studies where similarity was not manipulated (participants were not paired with a similar or dissimilar other), or they did not know the extent of similarity, the similarity-attraction hypothesis was not supported (Sprecher et al., 2015; Tidwell et al., 2013). Some CMC systems did not reveal the similarity manipulation directly but let participants find it through interaction (Kaptein et al., 2014; Ludford et al., 2004). Chatbuddies in Chapter 5 chose to reveal the similarity information and explanation the mechanism behind manipulation work for the transparency of the system. In research studies there is no problem to hide the similarity information for the sake of study purpose. But if an algorithm is applied in a system applied similarity-based algorithm, it is better to inform users.

Type and Stage of Friendship. Studies in Chapter 4 and Chapter 5 assumed users are strangers and are seeking friends. But it is also common that offline friends or users who already developed strong friendships join a mass interaction together. In this case the system should consider adding social network in the grouping algorithm.

Perceived or Actual Similarity. Though some studies claimed that it is perceived similarity instead of actual similarity that contributes to attraction (Montoya et al., 2008; Sprecher, 2014; Sunnafrank, 1983, 1984, 1985, 1986), as discussed in Chapter 2, given the complexity of similarity dimensions, the difficulty of accurately measuring similarity, and some hidden variables such as the importance of certain dimensions in relationship development, it seems arbitrary to reject the effect of actual similarity. Literature (Cosley et al., 2003; Ludford et al., 2004) and the lab study in Chapter 4 revealed that people can recognize actual similarity on certain dimensions by interacting with others. Instead of examining which one is the real cause, Chatbuddies in

Chapter 5 utilized the findings from previous literature and applied both actual and perceived similarity.

Ethical Issues

The issues of ethics have been discussed for a long time in both design (Borning & Muller, 2012; Friedman et al., 2013) and machine learning domains (Anderson & Anderson, 2011; Dietterich & Horvitz, 2015). Both domains advocate a thorough conceptual, empirical, and technical analysis to understand the short- and long-term risks of digital technologies.

Subgrouping in large-scale communication faces at least three ethical issues. First, is a live stream system granted permission to create subgroups? Nowadays many social media platforms apply machine learning algorithms for "better" user experience, such as news feed, friend recommendations, and romantic partner matching. Users must agree with some terms to get access to the service, including data collection and the use of algorithms (Flick, 2015). Is it ethical to manipulate the content a user interacts with for sake of their experience? Regulating social media algorithms remains lots of debates in laws and policies (Julia Zorthain, 2021) and is technically challenging (Cen & Shah, 2021). In response to users' resistance of algorithms (Devito et al., 2017) and increasing requests to disable algorithms (Ruvic, 2021), some platforms give users options to opt out of algorithmic recommendations, such as going back to a reverse-chronological newsfeed (Bell, 2022; Newton, 2018). Designers should respect users' preferences of not being sub-grouped.

Second, the process of subgrouping itself may bring ethical issues. In Chapter 4 and Chapter 5, I chose interests, common ground, and channel seniority as similarity dimensions because they are supported by empirical studies and because these features are accessible to developers. An underlying criterion for selecting similarity dimensions is ethical integrity and fairness. Machine learning fairness literature has pointed out the importance of fairness through unawareness (Gajane & Pechenizkiy, 2017; Sahil Verma & Julia Rubin, 2018), which argues that protective attributes such as race and gender should not be explicitly used in decision-making (Grgic-Hlaca et al., 2016). Designers and practitioners are supposed to carefully screen attributes that are not only effective but also ethical to use.

Third, personalized filtering algorithms often bring biases and may have negative consequences on both individuals and communities (Bozdag, 2013). Subgrouping can be seen as a filtering method because it filters some viewers' chat messages. As I mentioned above, similarity-based grouping may create echo chambers or filter bubbles (Pariser, 2011; Sunstein, 2018). Echo chambers are not inherently bad, in many situations it guarantees shared assumptions that help conversations to move on (Bruckman, 2022). But the polarity of social media ends up reaffirming people's existing positions instead of diversifying opinions (Nelimarkka et al., 2018). Similarity-based subgroups may be criticized for creating positive bias. Besides similarity-based methods, other algorithmic subgrouping methods may face the accusation of manipulation because research showed that chat messages can influence viewers' political voting and attitude (Maruyama et al., 2014, 2017), emotions and moods (Kramer et al., 2014), anti-social behaviors (J. Cheng et al., 2017), and gifting behaviors (R. Li et al., 2020). SNS platforms have a long history of experimenting with algorithms to reach different commercial goals, but critiques never stop (Merrill & Oremus, 2021; Newton, 2018; Oremus et al., 2021). Designers and practitioners may not intend to carry out the manipulation, but they should be vigilant about potential side effects of any subgrouping algorithm and have a good judgement on its ethical consequence. Brown, Davidovic, & Hasan (2021) proposed a framework to guide the ethical assessment of an algorithm: a list of the interests of stakeholders who may be affected by the algorithm, an assessment of metrics that describe ethically salient features of the algorithm, and a relevancy matrix that connects the metrics to stakeholder interests.

Design Future AI-embedded CMC

In Chapter 2, I introduce a concept of AI-embedded CMC that not only mediates the messages but also manipulate the communication structure. Recommendation system is one of the mostly studied AI-embedded CMC systems, including people or community recommendation and conversation recommendation. In this dissertation, I investigated AI-embedded CMC through live streaming, a type of real-time communication that integrates both video and text chat. In Chapter 5, I proposed a framework with three steps: detect subgroup need, generate subgroups, and visualize grouping results. As an intelligent system, it should have one more step that is to collect feedback and update the grouping algorithm. This dissertation used basic clustering models, therefore the system was not as "intelligent" as many commercial products. My goal is to propose a framework and evaluate its effects of human users. Future AI-embedded CMC system can leverage techniques in other areas such as natural language processing.

Thread Disentanglement is a task to figure out the "reply to" relationship among the utterances (Kummerfeld et al., 2020). Given a history of chat messages, if a system can predict the next message is responding to whom, it will help interlocutors to understand the conversation structure. As participants reported in Chapter 5, asymmetric subgrouping creates a unique subgroup for each viewer. As a result, if Viewer A see Viewer B's reply to Viewer C, who is not in Viewer A's subgroup, Viewer A may feel confused. With thread disentanglement, Viewer A can see the conversation thread between B and C. An overall conversation threads can be presented to both viewers and streamers.

Conversation Summarization can help newly joined viewers and streamers quickly get an overview of ongoing topics. Though chat summarization has not been thoroughly explored and has many difficulties because multi-party chat messages are shorter, less structured, and have more misspellings and acronyms (Uthus & Aha, 2013). It is an area worth investigation and can be combined with human annotation to bring both intellectual and communicative convenience (Lu, Heo, et al., 2018; A. X. Zhang & Cranshaw, 2018).

A Transparent and Explainable AI-embedded CMC System

A surge of interest in transparent and explainable AI has led to a large of body of research in computer science, HCI, and social science domains (Abdul et al., 2018; Arrieta et al., 2020; Long & Magerko, 2020; T. Miller, 2019; Shneiderman, 2020). None or insufficient explanation of a system's algorithms will lead to deception and misunderstanding (Eslami et al., 2019). Though Explainable AI (XAI) is a recent topic, HCI community has a long history of designing and evaluating explainable systems, such as recommendation systems (Cramer et al., 2008; Schaffer et al., 2015) , personalization (El-Arini et al., 2012), and grading (Kizilcec, 2016). It was reported that transparent explanations contribute to enhance acceptance of specific recommendations (Cramer et al., 2008) but diminish the user experience (Schaffer et al., 2015).

It is not the goal of this dissertation to examine the effectiveness of explanations in AI-embedded CMC systems. Chatbuddies applies design suggestions for AI systems from previous literature (Liao et al., 2020). The algorithm used in Chatbuddies is relatively simple, compared with some systems that embed deep neural networks trained on massive datasets. Therefore, the explanation is also straightforward. Chatbuddies provides both global explanation (e.g., "You are viewing chats from 52 viewers who are most similar to you") and local explanation (e.g., a profile page that shows channels and games follow in common, triggered by clicking the dot on the scatter plot or clicking the username in the chat window). Most participants in the online experiment did not report confusion about the explanations. But it took much effort to iterate the text and visualization design in the pilot phrase, as demonstrated in Chapter 5, which indicates that multiple iterations of explanation should be included in the system design.

General Limitations

I want to point out several limitations that should be considered when building on future research. First, though the research domain is large-scale communication, and this work focuses on live streaming context, the audience size of most live streams falls into a long-tail distribution, which means the majority of live streams have a small audience count (Seering et al., 2018). Therefore, the design and evaluation may not be applicable for all the live streams. Second, due to the recruitment difficulties discussed in Chapter 5, experiments in this dissertation did not involve real live streams where hundreds of participants join the chat room at the same time; instead, simulated live streams and deception were employed to achieve a similar research goal. The lack of real interaction may affect the ecological validity of the results. Third, most watching sessions in my studies are short (around 15 minutes). Long-term effects of similarity-baesd subgrouping were not investigated. Fourth, this dissertation focuses mostly on viewers' experiences as they are the majority participants of large-scale communication. But the streamer also has an important role in live streams. Future research should take streamers' need into account and design tools from their perspectives. Fifth, in this dissertation I examined live streams in domains such as gaming, IRL (in real life), and talk shows, but there are much more categories such as concerts, political debates, education. The results presented above may not apply in certain domains, even may bring negative consequences. Last, I experimented with multiple clustering algorithms and applied one of them in an online experiment. The study used simulated "live" streams, and viewers' similarity scores were precomputed, therefore participants can get grouped almost in real time. However, in real-world creating subgroups for a live stream with millions of viewers may take a long time, depending on the computation resource a live stream platform has.

Conclusion and Future Directions

This dissertation presents a series of work on understanding practices and challenges in large-scale real-time interaction and how to design new CMC tools to facilitate communication and create opportunities of social network building. Live streaming as a relatively new medium and social media feature was selected as the research context. From the user research, I summarized the motivations, understood common practices, and identified challenges of viewers when engaging with live streams on SNS. I proposed a framework that applies similarity-attraction effect in social psychology to large-scale live streams. Building on the results of a preliminary controlled lab experiment, I designed and developed Chatbuddies, an intelligent interface that generates subgroups based on similarity of common grounds and community seniority, utilizing algorithmic approaches. I proposed AI-embedded CMC, in which AI not only mediate message content but also communication structure. My work contributes to the application of social psychology theories on CMC applications and opens new directions of AI-embedded CMC.

REFERENCE

Abdul, A., Vermeulen, J., Wang, D., Lim, B. Y., & Kankanhalli, M. (2018). Trends and trajectories for explainable, accountable and intelligible systems: an hci research agenda. *Proceedings Of Conference On Human Factors In Computing Systems (CHI'18)*, *2018-April*.

Andalibi, N., & Flood, M. K. (2021). Considerations in designing digital peer support for mental health: interview study among users of a digital support system (buddy project). *JMIR Mental Health*, *8*(1), e21819.

Anderson, M., & Anderson, S. L. (Eds.). (2011). *Machine Ethics*. Cambridge University Press.

Aron, A., Melinat, E., Aron, E. N., Vallone, R. D., & Bator, R. J. (1997). The experimental generation of interpersonal closeness: a procedure and some preliminary findings: *Personality And Social Psychology Bulletin*, *23*(4), 363–377.

Arrieta, A. B., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., Garcia, S., Gil-Lopez, S., Molina, D., Benjamins, R., Chatila, R., & Herrera, F. (2020). Explainable artificial intelligence (xai): concepts, taxonomies, opportunities and challenges toward responsible ai. *Information Fusion*, *58*, 82–115.

Bajarin, T. (2021, April 29). Work from home is the new normal for workers around the world. *Forbes*.

Banikiotes, P. G., & Neimeyer, G. J. (1981). Construct importance and rating similarity as determinants of interpersonal attraction. *British Journal Of Social Psychology*, *20*(4), 259–263.

Baumeister, R. F., & Vohs, K. D. (2007). *Encyclopedia of social psychology*. SAGE Publications, Inc.

Baxter, L. A., & West, L. (2003). Couple perceptions of their similarities and differences: a dialectical perspective: *Journal Of Social And Personal Relationships*, *20*(4), 491–514.

Bell, K. (2017, September 1). Facebook tests feature that makes it easier to creep on mutual friends. *Mashable*.

Bell, K. (2022, March 23). *Instagram's chronological feed is back*. Engadget.

Berscheid, E., & Reis, H. T. (1998). Attraction and close relationships. In D. T. Gilbert, S. T. Fiske, & Lindzey G. (Eds.), *The Handbook of Social Psychology* (pp. 193–281).

Bertini, M., Ferracani, A., Papucci, R., & Del Bimbo, A. (2020). Keeping up with the influencers: improving user recommendation in instagram using visual content. *UMAP 2020 Adjunct - Adjunct Publication Of The 28th ACM Conference On User Modeling, Adaptation And Personalization*, 29–34.

Borning, A., & Muller, M. (2012). Next steps for value sensitive design. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 1125–1134). ACM.

Boy, G. A. (2011). *The Handbook of Human-Machine Interaction: A Human-Centered Design Approach*. CRC Press.

Boyd, D. M., & Ellison, N. B. (2007). Social network sites: definition, history, and scholarship. *Journal Of Computer-Mediated Communication*, *13*(1), 210–230.

Bozdag, E. (2013). Bias in algorithmic filtering and personalization. *Ethics And Information Technology*, *15*(3), 209–227.

Brewer, R. E., & Brewer, M. B. (1968). Attraction and accuracy of perception in dyads. *Journal Of Personality And Social Psychology*, *8*(2 PART 1), 188–193.

Brown, S., Davidovic, J., & Hasan, A. (2021). The algorithm audit: scoring the algorithms that score us. *Big Data & Society*, *8*(1).

Bruckman, A. S. (2022). *Should You Believe Wikipedia?: Online Communities and the*

Construction of Knowledge. Cambridge University Press.

Buckels, E. E., Trapnell, P. D., & Paulhus, D. L. (2014). Trolls just want to have fun. *Personality And Individual Differences, 67*, 97–102.

Byrne, D. (1961). Interpersonal attraction and attitude similarity. *Journal Of Abnormal And Social Psychology*, *62*(3), 713–715.

Byrne, D. (1971). The attraction paradigm. In *Academic Press*.

Byrne, D. (1992). The transition from controlled laboratory experimentation to less controlled settings: surprise! additional variables are operative. *Communication Monographs* , *59*(2), 190–198.

Byrne, D., & Clore, G. L. (1966). Predicting interpersonal attraction toward strangers presented in three different stimulus modes. *Psychonomic Science*, *4*(6), 239–240.

Byrne, D., Clore, G. L., & Worchel, P. (1966). Effect of economic similarity-dissimilarity on interpersonal attraction. *Journal Of Personality And Social Psychology*, *4*(2), 220–224.

Byrne, D., Ervin, C. R., & Lamberth, J. (1970). Continuity between the experimental study of attraction and real-life computer dating. *Journal Of Personality And Social Psychology, 16*(1), 157–165.

Byrne, D., Gouaux, C., Griffitt, W., Lamberth, J., Murakawa, N., Prasad, M. B., Prasad, A., & Ramirez, M. (1971). The ubiquitous relationship: attitude similarity and attraction: a cross-cultural study. *Human Relations, 24*(3), 201–207.

Byrne, D., & Griffitt, W. (1966). A developmental investigation of the law of attraction. *Journal Of Personality And Social Psychology*, *4*(6), 699–702.

Byrne, D., Griffitt, W., & Stefaniak, D. (1967). Attraction and similarity of personality characteristics. *Journal Of Personality And Social Psychology*, *5*(1), 82–90.

Byrne, D., Hudgins, W., Reeves, K., & Griffitt, W. (1969). Attitude similarity-dissimilarity and attraction: generality beyond the college sophomore. *The Journal Of Social Psychology*, *79*(2), 155–161.

Byrne, D., & Nelson, D. (1965). Attraction as a linear function of proportion of positive reinforcements. *Journal Of Personality And Social Psychology*, *1*(6), 659–663.

Cai, J., Wohn, Y., Mittal, A., & Sureshbabu, D. (2018). Utilitarian and hedonic motivations for live streaming shopping. *Proceedings Of The 2018 ACM International Conference On Interactive Experiences For TV And Online Video*, 81–88.

Cameron, C. D., Brown-Iannuzzi, J. L., & Payne, B. K. (2012). Sequential priming measures of implicit social cognition: a meta-analysis of associations with behavior and explicit attitudes. *Personality And Social Psychology Review*, *16*(4), 330–350.

Cappella, J. N., & Palmer, M. T. (1990). Attitude similarity, relational history, and attraction: the mediating effects of kinesic and vocal behaviors. *Communication Monographs*, *57*(3), 161–183.

Carter, D. M. (2004). Living in virtual communities: making friends online. *Journal Of Urban Technology* , *11*(3), 109–125.

Cen, S., & Shah, D. (2021). Regulating algorithmic filtering on social media. *Advances In Neural Information Processing Systems 34 (NeurIPS 2021)*, 6997–7011.

Chan, D. K. S., & Cheng, G. H. L. (2004). A comparison of offline and online friendship qualities at different stages of relationship development: *Journal Of Social And Personal Relationships*, *21*(3), 305–320.

Chen, J., Geyer, W., Dugan, C., Muller, M., & Guy, I. (2009). Make new friends, but keep the old: recommending people on social networking sites. *Proceedings Of The 27th International*

Conference On Human Factors In Computing Systems (CHI'09), 201–210.

Chen, J., Nairn, R., & Chi, E. H. (2011). Speak little and well: recommending conversations in online social streams. *Proceedings Of The SIGCHI Conference On Human Factors In Computing Systems (CHI'11)*, 217–226.

Chen, Y., Lasecki, W. S., & Dong, T. (2020). Towards supporting programming education at scale via live streaming. *Proceedings Of The ACM On Human-Computer Interaction*, *4*(CSCW3), 1–19.

Chen, Y., Lee, S. W., Xie, Y., Yang, Y., Lasecki, W. S., & Oney, S. (2017). Codeon: on-demand software development assistance. *Proceedings Of The 2017 CHI Conference On Human Factors In Computing Systems (CHI'17)*, 6220–6231.

Cheng, H.-F., Wang, R., Zhang, Z., O'Connell, F., Gray, T., Harper, F. M., & Zhu, H. (2019). Explaining decision-making algorithms through ui: strategies to help non-expert stakeholders. *Proceedings Of The 2019 CHI Conference On Human Factors In Computing Systems (CHI '19)*, 1–12.

Cheng, J., Bernstein, M., Danescu-Niculescu-Mizil, C., & Leskovec, J. (2017). Anyone can become a troll: causes of trolling behavior in online discussions. *Proceedings Of The 2017 ACM Conference On Computer Supported Cooperative Work And Social Computing (CSCW '17)*, 1217–1230.

Civan, A., McDonald, D. W., Unruh, K. T., & Pratt, W. (2009). Locating patient expertise in everyday life. *Proceedings Of The 2009 ACM SIGCHI International Conference On Supporting Group Work (GROUP'09)*, 291–300.

Clark, H. H., & Brennan, S. E. (1991). Grounding in communication. In *Perspectives on socially shared cognition* (pp. 127–149).

Clore, G. L., & Byrne, D. (1974). A reinforcement-affect model of attraction. *Foundations Of Interpersonal Attraction*, 143–170.

Cooley, M. (2000). Human-centered design. In R. Jacobson (Ed.), *Information Design* (pp. 59–81). The MIT Press.

Corbin, J., & Strauss, A. (2012). Basics of qualitative research (3rd ed.): techniques and procedures for developing grounded theory. In *Basics of Qualitative Research (3rd ed.): Techniques and Procedures for Developing Grounded Theory*.

Cosley, D., Ludford, P., & Terveen, L. (2003). Studying the effect of similarity in online task-focused interactions. *Proceedings Of The 2003 International ACM SIGGROUP Conference On Supporting Group Work (GROUP '03)*, 321–329.

Craker, N., & March, E. (2016). The dark side of facebook: the dark tetrad, negative social potency, and trolling behaviours. *Personality And Individual Differences*, *102*, 79–84.

Cramer, H., Evers, V., Ramlal, S., Maarten Van Someren, ·, Rutledge, L., Stash, N., Aroyo, L., Wielinga, B., Cramer, H., Evers, V., Ramlal, · S, Van Someren, · M, Ramlal, S., Van Someren, M., Rutledge, L., Aroyo, L., Stash, N., Aroyo, · L, & Wielinga, B. (2008). The effects of transparency on trust in and acceptance of a content-based art recommender. *User Modeling And User-Adapted Interaction 2008 18:5*, *18*(5), 455–496.

Curry, T. J., & Emerson, R. M. (1970). Balance theory: a theory of interpersonal attraction? *Sociometry, 33*(2), 216–238.

Dahimene, R., Constantin, C., & Du Mouza, C. (2014). RecLand: a recommender system for social networks. *CIKM 2014 - Proceedings Of The 2014 ACM International Conference On Information And Knowledge Management*, 2063–2065.

de Oliveira, R., Pentoney, C., & Pritchard-Berman, M. (2018). YouTube needs: understanding user's motivations to watch videos on mobile devices. *Proceedings Of The 20th International Conference On Human-Computer Interaction With Mobile Devices And Services (MobileHCI '18)*, 1–11.

Deshpande, A., Kacham, P., & Pratap, R. (2020). Robust k-means++. *Conference On Uncertainty In Artificial Intelligence.*, *124*, 799–808.

Devito, M. A., Gergle, D., & Birnholtz, J. (2017). "Algorithms ruin everything": #riptwitter, folk theories, and resistance to algorithmic change in social media. *Proceedings Of The 2017 CHI Conference On Human Factors In Computing Systems (CHI'17)*, 3163–3174.

Dietterich, T. G., & Horvitz, E. J. (2015). Rise of concerns about ai: reflections and directions. *Communications Of The ACM*, *58*(10), 38–40.

Ding, X., Jin, X., Li, Y., & Li, L. (2013). Celebrity recommendation with collaborative social topic regression. *International Joint Conference On Artificial Intelligence (IJCAI '13)*, 2612–2618.

Douglas, K. M., & McGarty, C. (2001). Identifiability and self-presentation: computer-mediated communication and intergroup interaction. *British Journal Of Social Psychology*, *40*(3), 399–416.

Duck, S. W., & Craig, G. (1978). Personality similarity and the development of friendship: a longitudinal study. *British Journal Of Social And Clinical Psychology*, *17*(3), 237–242.

El-Arini, K., Paquet, U., Herbrich, R., Van Gael, J., & Agüera Y Arcas, B. (2012). Transparent user models for personalization. *Proceedings Of The ACM SIGKDD International Conference On Knowledge Discovery And Data Mining*, 678–686.

Eslami, M., Vaccaro, K., Lee, M. K., Elazari, A., On, B., Gilbert, E., Karahalios, K., & Elazari Bar On, A. (2019). User attitudes towards algorithmic opacity and transparency in online reviewing platforms. *Proceedings Of The 2019 CHI Conference On Human Factors In*

Computing Systems (CHI'19), 1–14.

Ester, M., Kriegel, H.-P., Sander, J., & Xu, X. (1996). A density-based algorithm for discovering clusters in large spatial databases with noise. *Proceedings Of The Conference On Knowledge Discovery & Data Mining*, 2226–2231.

Faas, T., Dombrowski, L., Young, A., & Miller, A. D. (2018). Watch me code: programming mentorship communities on twitch.tv. *Proceedings Of The ACM On Human-Computer Interaction*, *2*(CSCW), 1–18.

Facebook. (2020). *Insights to go: Search for insights from Facebook IQ*. Facebook.

Fan, C., Hao, H., Leung, C. K., Sun, L. Y., & Tran, J. (2018). Social network mining for recommendation of friends based on music interests. *Proceedings Of The 2018 IEEE/ACM International Conference On Advances In Social Networks Analysis And Mining, ASONAM 2018*, 833–840.

Farzan, R., Dabbish, L., Kraut, R., & Postmes, T. (2011). Increasing commitment to online communities by designing for social presence. *Proceedings Of The ACM 2011 Conference On Computer Supported Cooperative Work (CSCW '11)*, 321–330.

Filippone, M., Camastra, F., Masulli, F., & Rovetta, S. (2008). A survey of kernel and spectral methods for clustering. *Pattern Recognition*, *41*(1), 176–190.

Finkel, E. J., Eastwick, P. W., Karney, B. R., Reis, H. T., & Sprecher, S. (2012). Online dating: a critical analysis from the perspective of psychological science. *Psychological Science In The Public Interest*, *13*(1), 3–66.

Flick, C. (2015). Informed consent and the facebook emotional manipulation study: *Research Ethics*, *12*(1), 14–28.

Foramitti, J., Drews, S., Klein, F., & Konc, T. (2021). The virtues of virtual conferences. *Journal*

Of Cleaner Production, *294*, 126287.

Ford, C., Gardner, D., Horgan, L. E., Liu, C., Tsaasan, a. m., Nardi, B., & Rickman, J. (2017). Chat speed op pogchamp: practices of coherence in massive twitch chat. *Proceedings Of The 2017 CHI Conference Extended Abstracts On Human Factors In Computing Systems (CHI EA '17)*, 858–871.

Fraser, C. A., Kim, J. O., Thornsberry, A., Klemmer, S., & Dontcheva, M. (2019). Sharing the studio: how creative livestreaming can inspire, educate, and engage. *Proceedings Of The 2019 On Creativity And Cognition (C&C '19)*, 144–155.

Friedman, B., Kahn, P. H., Borning, A., & Huldtgren, A. (2013). Value sensitive design and information systems. *Philosophy Of Engineering And Technology*, *16*, 55–95.

Gajane, P., & Pechenizkiy, M. (2017). *On Formalizing Fairness in Prediction with Machine Learning*.

Gil de Zúñiga, H., Garcia-Perdomo, V., & McGregor, S. C. (2015). What is second screening? exploring motivations of second screen use and its effect on online political participation. *Journal Of Communication*, *65*(5), 793–815.

Gómez-Zará, D., Guo, M., Dechurch, L. A., & Contractor, N. (2020). The impact of displaying diversity information on the formation of self-assembling teams. *Proceedings Of Conference On Human Factors In Computing Systems (CHI'20)*, 1–15.

Good, L. R., & Good, K. C. (1974a). Similarity of attitudes and attraction to an occupation. *Psychological Reports*, *35*(2), 703–706.

Good, L. R., & Good, K. C. (1974b). Similarity of attitudes and attraction to a social organization. *Psychological Reports*, *34*(3_suppl), 1071–1073.

Grgic-Hlaca, N., Zafar, M. B., Gummadi, K. P., & Weller, A. (2016). The case for process fairness in learning: feature selection for fair decision making. *NIPS Symposium On Machine Learning And The Law*.

Griffitt, W., & Veitch, R. (1974). Preacquaintance attitude similarity and attraction revisited: ten days in a fall-out shelter. *Sociometry, 37*(2), 163.

Guo, D., Xu, J., Zhang, J., Xu, M., Cui, Y., & He, X. (2017). User relationship strength modeling for friend recommendation on instagram. *Neurocomputing*, *239*, 9–18.

Guo, J., & Fussell, S. R. (2020). A preliminary study of emotional contagion in live streaming. *Conference Companion Publication Of The 2020 On Computer Supported Cooperative Work And Social Computing*, 263–268.

Guy, I. (2018). People recommendation on social media. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics): Vol. 10100 LNCS* (pp. 570–623). Springer Verlag.

Guy, I., Ur, S., Ronen, I., Perer, A., & Jacovi, M. (2011). Do you want to know?: recommending strangers in the enterprise. *Proceedings Of The ACM Conference On Computer Supported Cooperative Work, CSCW*, 285–294.

Haimson, O. L., & Tang, J. C. (2017). What makes live events engaging on facebook live, periscope, and snapchat. *Proceedings Of The 2017 CHI Conference On Human Factors In Computing Systems (CHI '17)*, 48–60.

Hamilton, W. A., Garretson, O., & Kerne, A. (2014). Streaming on twitch: fostering participatory communities of play within live mixed media. *Proceedings Of The SIGCHI Conference On Human Factors In Computing Systems (CHI '14)*, 1315–1324.

Hampton, A. J., Boyd, A. N. F., & Sprecher, S. (2018). You're like me and i like you: mediators of the similarity–liking link assessed before and after a getting-acquainted social interaction:

Journal Of Social And Personal Relationships, *36*(7), 2221–2244.

Hancock, J., Naaman, M., & Levy, K. (2020). AI-mediated communication: definition, research agenda, and ethical considerations. *Journal Of Computer-Mediated Communication*, *25*(1), 89–100.

Haner, D., & Pepler, D. (2016). “Live chat” clients at kids help phone: individual characteristics and problem topics. *Journal Of The Canadian Academy Of Child And Adolescent Psychiatry*, *25*(3), 138.

Hansson, R. O., & Fiedler, F. E. (1973). Perceived similarity, personality, and attraction to large organizations. *Journal Of Applied Social Psychology*, *3*(3), 258–266.

Harboe, G., Metcalf, C. J., Bentley, F., Tullio, J., Massey, N., & Romano, G. (2008). Ambient social tv. *Proceedings Of The SIGCHI Conference On Human Factors In Computing Systems (CHI ’08)*, 1–10.

Harrison, D. A., Price, K. H., Gavin, J. H., & Florey, A. T. (2002). Time, teams, and task performance: changing effects of surface- and deep-level diversity on group functioning. *Academy Of Management Journal*, *45*(5), 1029–1045.

Hart, S. G., & Staveland, L. E. (1988). Development of nasa-tlx (task load index): results of empirical and theoretical research. In P. A. Hancock & N. Meshkati (Eds.), *Human Mental Workload* (Vol. 52, pp. 139–183). North-Holland.

Hartzler, A. L., & Pratt, W. (2011). Managing the personal side of health: how patient expertise differs from the expertise of clinicians. *Journal Of Medical Internet Research*, *13*(3), e1728.

Hartzler, A. L., Taylor, M. N., Park, A., Griffiths, T., Backonja, U., McDonald, D. W., Wahbeh, S., Brown, C., & Pratt, W. (2016). Leveraging cues from person-generated health data for peer matching in online communities. *Journal Of The American Medical Informatics*

Association, *23*(3), 496–507.

Hauben, M., & Hauben, R. (1997). *Netizens : on the history and impact of Usenet and the Internet*. IEEE Computer Society Press.

Helsen, K., Derom, I., Corthouts, J., Bosscher, V. De, Willem, A., & Scheerder, J. (2021). Participatory sport events in times of covid-19: analysing the (virtual) sport behaviour of event participants. *European Sport Management Quarterly*, *22*(1), 35–54.

Herring, S. C. (1999). Interactional coherence in cmc. *Proceedings Of The 32nd Annual Hawaii International Conference On Systems Sciences. 1999. HICSS-32. Abstracts And CD-ROM Of Full Papers*, 1–13.

Herring, S. C. (2002). Computer-mediated communication on the internet. *Annual Review Of Information Science And Technology*, *36*(1), 109–168.

Hester, L. R., Weitz, L. J., Anchor, K. N., & Roback, H. B. (1976). Supervisor attraction as a function of level of supervisor skillfulness and supervisees' perceived similarity. *Journal Of Counseling Psychology*, *23*(3), 254–258.

Hilvert-Bruce, Z., Neill, J. T., Sjöblom, M., & Hamari, J. (2018). Social motivations of live-streaming viewer engagement on twitch. *Computers In Human Behavior*, *84*, 58–67.

Hogg, M. A. (1993). Group cohesiveness: a critical review and some new directions. *European Review Of Social Psychology*, *4*(1), 85–111.

Hohenstein, J., & Jung, M. (2018). AI-supported messaging: an investigation of human-human text conversation with ai support. *Extended Abstracts Of The 2018 CHI Conference On Human Factors In Computing Systems (CHI EA '18)*, 1–6.

Holtzblatt, K., & Beyer, H. (1997). *Contextual Design-Defining Customer-Centred Systems*. Elsevier.

Huston, T. L., & Levinger, G. (1978). Interpersonal attraction and relationships. *Annual Review Of Psychology*, *29*, 115–156.

IBM. (2020). *Clustering binary data with K-Means (should be avoided)*. IBM Support.

Instagram. (2018). *How do I view someone's live video? | Instagram Help Centre*.

Isaacs, E., Szymanski, M., Yamauchi, Y., Glasnapp, J., & Iwamoto, K. (2012). Integrating local and remote worlds through channel blending. *Proceedings Of The ACM 2012 Conference On Computer Supported Cooperative Work (CSCW '12)*, 617–626.

Joinson, A. N. (2008). "Looking at", "Looking up" or "keeping up with" people? motives and uses of facebook. *Proceedings Of The SIGCHI Conference On Human Factors In Computing Systems (CHI '08)*, 1027–1036.

Jones, Q., Moldovan, M., Raban, D. R., & Butler, B. S. (2008). Empirical evidence of information overload constraining chat channel community interactions. *Proceedings Of The ACM 2008 Conference On Computer Supported Cooperative Work (CSCW '08)*, 323.

Jones, Q., Ravid, G., & Rafaeli, S. (2004). Information overload and the message dynamics of online interaction spaces: a theoretical model and empirical exploration. *Information Systems Research*, *15*(2), 194–211.

Julia Zorthain. (2021, October 13). Why regulating facebook's algorithm is legally challenging | time. *TIME*.

Kahneman, D. (1973). *Attention and effort* (Vol. 1063). Englewood Cliffs.

Kaptein, M., Castaneda, D., Fernandez, N., & Nass, C. (2014). Extending the similarity-attraction effect: the effects of when-similarity in computer-mediated communication. *Journal Of Computer-Mediated Communication*, *19*(3), 342–357.

Karimpour, D., Zare Chahooki, M. A., & Hashemi, A. (2021). User recommendation based on hybrid filtering in telegram messenger. *26th International Computer Conference, Computer Society Of Iran, CSICC 2021*, 1–7.

Kaytoue, M., Silva, A., Cerf, L., Meira, W., & Raïssi, C. (2012). Watch me playing, i am a professional: a first study on video game live streaming. *Proceedings Of The 21st International Conference Companion On World Wide Web (WWW '12 Companion)*, 1181–1188.

Kim, J. (2013). Influence of group size on students' participation in online discussion forums. *Computers & Education*, *62*, 123–129.

Kizilcec, R. F. (2016). How much information? effects of transparency on trust in an algorithmic interface. *Proceedings Of The 2016 CHI Conference On Human Factors In Computing Systems*, 2390–2395.

Koroleva, K., & Bolufé Röhler, A. J. (2012). Reducing information overload: design and evaluation of filtering & ranking algorithms for social networking sites. *ECIS 2012 Proceedings*.

Kramer, A. D. I. I., Guillory, J. E., & Hancock, J. T. (2014). Experimental evidence of massive-scale emotional contagion through social networks. *Proceedings Of The National Academy Of Sciences Of The United States Of America*, *111*(24), 8788–8790.

Kraut, R. E., & Resnick, P. (2012). *Building Successful Online Communities*. MIT Press.

Kumar, N., Yadav Mudda, K., Trishal, G., Konjengbam, A., & Singh, M. (2019). Where to post: routing questions to right community in community question answering systems. *Proceedings Of The ACM India Joint International Conference On Data Science And Management Of Data*, 136–142.

Kummerfeld, J. K., Gouravajhala, S. R., Peper, J. J., Athreya, V., Gunasekara, C., Ganhotra, J., Patel, S. S., Polymenakos, L., & Lasecki, W. S. (2020). A large-scale corpus for conversation disentanglement. *Proceedings Of The 57th Annual Meeting Of The Association For Computational Linguistics*, 3846–3856.

Lampe, C., Ellison, N., & Steinfield, C. (2007). A familiar face(book): profile elements as signals in an online social network. *Proceedings Of The SIGCHI Conference On Human Factors In Computing Systems*, 435–444.

Lang, A. (2000). The limited capacity model of mediated message processing. *Journal Of Communication*, *50*(1), 46–70.

Latzko-Toth, G. (2010). Metaphors of synchrony: emergence and differentiation of online chat devices: *Bulletin Of Science, Technology & Society*, *30*(5), 362–374.

Lee, Y.-C., Yen, C.-H., Chiu, P.-T., King, J.-T., & Fu, W.-T. (2018). Tip me!: tipping is changing social interactions on live streams in china. *Extended Abstracts Of The 2018 CHI Conference On Human Factors In Computing Systems - CHI '18*, 1–6.

Levinger, G. (1972). Little sand box and big quarry: comment on byrne's paradigmatic spade for research on interpersonal attraction. *Representative Research In Social Psychology*, *3*(1), 3–19.

Li, J., Gui, X., Kou, Y., & Li, Y. (2019). Live streaming as co-performance: dynamics between center and periphery in theatrical engagement. *Proceedings Of The ACM On Human-Computer Interaction*, *3*(CSCW), 1–22.

Li, R., Lu, Y., Ma, J., & Wang, W. (2020). Examining gifting behavior on live streaming platforms: an identity-based motivation model. *Information And Management*, *58*(6), 103406.

Liao, Q. V., Gruen, D., & Miller, S. (2020). Questioning the ai: informing design practices for explainable ai user experiences. *Proceedings Of The 2020 CHI Conference On Human*

Factors In Computing Systems, 1–15.

Long, D., & Magerko, B. (2020). What is ai literacy? competencies and design considerations. *Proceedings Of The 2020 CHI Conference On Human Factors In Computing Systems*, 1–16.

Lorenz, T. (2021, March 18). The endless stream. *The New York Times*.

Lu, Z., Annett, M., Fan, M., & Wigdor, D. (2019). "I feel it is my responsibility to stream": streaming and engaging with intangible cultural heritage through livestreaming. *Proceedings Of The 2019 CHI Conference On Human Factors In Computing Systems*, 1–14.

Lu, Z., Annett, M., & Wigdor, D. (2019). Vicariously experiencing it all without going outside. *Proceedings Of The ACM On Human-Computer Interaction*, *3*(CSCW), 1–28.

Lu, Z., Heo, S., & Wigdor, D. J. (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), 1–26.

Lu, Z., Xia, H., Heo, S., & Wigdor, D. (2018). You watch, you give, and you engage: a study of live streaming practices in china. *Proceedings Of The 2018 CHI Conference On Human Factors In Computing Systems*, 1–13.

Ludford, P. J., Cosley, D., Frankowski, D., & Terveen, L. (2004). Think different: increasing online community participation using uniqueness and group dissimilarity. *Proceedings Of The 2004 Conference On Human Factors In Computing Systems (CHI '04)*, 631–638.

Luo, M., Hsu, T. W., Park, J. S., & Hancock, J. T. (2020). Emotional amplification during live-streaming: evidence from comments during and after news events. *Proceedings Of The ACM On Human-Computer Interaction*, *4*(CSCW1), 1–19.

Maloney-Krichmar, D., & Preece, J. (2005). A multilevel analysis of sociability, usability, and community dynamics in an online health community. *ACM Transactions On Computer-*

Human Interaction, *12*(2), 201–232.

Mamykina, L., Miller, A. D., Medynskiy, Y., Grevet, C., Davidson, P. R., Terry, M. A., Mynatt, E. D., & Cheriton, D. R. (2011). Examining the impact of collaborative tagging on sensemaking in nutrition management. *Proceedings Of The SIGCHI Conference On Human Factors In Computing Systems*, 657–666.

Maruyama, M., Robertson, S. P., Douglas, S., Raine, R., & Semaan, B. (2017). "Social watching" a civic broadcast: understanding the effects of positive feedback and other users' opinions. *Proceedings Of The ACM Conference On Computer Supported Cooperative Work, CSCW*, 794–807.

Maruyama, M., Robertson, S. P., Douglas, S., Semaan, B., & Faucett, H. (2014). Hybrid media consumption: how tweeting during a televised political debate influences the vote decision. *Proceedings Of The ACM Conference On Computer Supported Cooperative Work, CSCW*, 1422–1432.

McCarthy, B., & Duck, S. W. (1976). Friendship duration and responses to attitudinal agreement-disagreement. *British Journal Of Social And Clinical Psychology*, *15*(4), 377–386.

McCroskey, L. L., McCroskey, J. C., & Richmond, V. P. (2006). Analysis and improvement of the measurement of interpersonal attraction and homophily. *Communication Quarterly*, *54*(1), 1–31.

Mcdonald, D. W. (2003). Recommending collaboration with social networks: a comparative evaluation. *Proceedings Of The Conference On Human Factors In Computing Systems (CHI '03)*.

McInnis, B. J., Murnane, E. L., Epstein, D., Cosley, D., & Leshed, G. (2016). One and done: factors affecting one-time contributors to ad-hoc online communities. *Proceedings Of The*

19th ACM Conference On Computer-Supported Cooperative Work & Social Computing - CSCW '16, 608–622.

Merrill, J. B., & Oremus, W. (2021, October 26). Five points for anger, one for a 'like': how facebook's formula fostered rage and misinformation. *The Washington Post*.

Meta. (n.d.). *Facebook Live*. Meta for Media.

Michinov, E., & Michinov, N. (2011). Social comparison orientation moderates the effects of group membership on the similarity-attraction relationship. *The Journal Of Social Psychology*, *151*(6), 754–766.

Miller, M. K., Tang, J. C., Venolia, G., Wilkinson, G., & Inkpen, K. (2017). Conversational chat circles: being all here without having to hear it all. *Proceedings Of The 2017 CHI Conference On Human Factors In Computing Systems (CHI '17)*, 2394–2404.

Miller, T. (2019). Explanation in artificial intelligence: insights from the social sciences. *Artificial Intelligence*, *267*, 1–38.

Mohammed, M., Khan, M. B., & Bashie, E. B. M. (2016). *Machine learning: Algorithms and applications*. CRC Press.

Montoya, R. M., & Horton, R. S. (2012). A meta-analytic investigation of the processes underlying the similarity-attraction effect. *Journal Of Social And Personal Relationships*, *30*(1), 64–94.

Montoya, R. M., Horton, R. S., & Kirchner, J. (2008). Is actual similarity necessary for attraction? a meta-analysis of actual and perceived similarity. *Journal Of Social And Personal Relationships*, *25*(6), 889–922.

Musabirov, I., Bulygin, D., Okopny, P., & Konstantinova, K. (2018, April 20). Event-driven spectators' communication in massive esports online chats. *Extended Abstracts Of The 2018 CHI Conference On Human Factors In Computing Systems*.

Neimeyer, R. A., & Mitchell, K. A. (1988). Similarity and attraction: a longitudinal study: *Journal Of Social And Personal Relationships*, *5*(2), 131–148.

Neimeyer, R. A., & Neimeyer, G. J. (1983). Structural similarity in the acquaintance process. *Journal Of Social And Clinical Psychology*, *1*(2), 146–154.

Nelimarkka, M., Laaksonen, S. M., & Semaan, B. (2018). Social media is polarized, social media is polarized: towards a new design agenda for mitigating polarization. *Proceedings Of The 2018 On Designing Interactive Systems Conference 2018 (DIS '18)*, 957–970.

Nematzadeh, A., Ciampaglia, G. L., Ahn, Y. Y., & Flammini, A. (2019). Information overload in group communication: from conversation to cacophony in the twitch chat. *Royal Society Open Science*, *6*(10).

Newcomb, T. M. (1961). *The Acquaintance Process*. Holt, Rinehart & Winston.

Newton, C. (2016). Here's how twitter's new algorithmic timeline is going to work. *The Verge*.

Newton, C. (2018, December). Twitter is relaunching the reverse-chronological feed as an option for all users starting today. *The Verge*.

Newton, C. (2021, February 1). Elon musk just showed how clubhouse can succeed. *The Verge*.

Nielsen. (2020). *Nearly 21 Million Viewers Watch One World: Together at Home Concert*. The Nielsen Company.

Nieva, R. (2018, April 24). Facebook tests "things in common" label to try to connect non-friends. *CNET*.

O'Brien, H. L., & Toms, E. G. (2008). What is user engagement? a conceptual framework for defining user engagement with technology. *Journal Of The American Society For Information Science And Technology*, *59*(6), 938–955.

O'Brien, H. L., & Toms, E. G. (2010). The development and evaluation of a survey to measure user engagement. *Journal Of The American Society For Information Science And Technology*, *61*(1), 50–69.

O'Leary, K., Bhattacharya, A., Munson, S. A., Wobbrock, J. O., & Pratt, W. (2017). Design opportunities for mental health peer support technologies. *Proceedings Of The ACM Conference On Computer Supported Cooperative Work (CSCW'17)*, 1470–1484.

Oremus, W., Alcantara, C., Merrill, J. B., & Galocha, A. (2021, October 26). How facebook shapes your feed. *The Washington Post*.

Pariser, E. (2011). *The Filter Bubble: How the New Personalized Web Is Changing What We Read and How We Think*. Penguin.

Petrocelli, B. (2017). Your chat has been upgraded with followers-only mode. In *Twitch*.

Pfaffenberger, B. (2003). *"A Standing Wave in the Web of Our Communications": Usenet and the Socio-Technical Construction of Cyberspace Values*. 20–43.

Pinto, H., Almeida, J. M., & Gonçalves, M. A. (2013). Using early view patterns to predict the popularity of youtube videos. *WSDM 2013 - Proceedings Of The 6th ACM International Conference On Web Search And Data Mining*, 365–374.

Pizzato, L., Rej, T., Akehurst, J., Koprinska, I., Yacef, K., Kay, J., Pizzato, L., Rej, · T, Akehurst, · J, Koprinska, · I, Yacef, · K, Kay, · J, & Kay, J. (2012). Recommending people to people: the nature of reciprocal recommenders with a case study in online dating. *User Modeling And User-Adapted Interaction 2012 23:5*, *23*(5), 447–488.

Preece, J., Nonnecke, B., & Andrews, D. (2004). The top five reasons for lurking: improving community experiences for everyone. *Computers In Human Behavior*, *20*(2), 201–223.

Rai, P., & Singh, S. (2010). A survey of clustering techniques. *International Journal Of Computer Applications*, *7*(12), 1–5.

Raman, A., Tyson, G., & Sastry, N. (2018). Facebook (a)live? are live social broadcasts really broadcasts? *Proceedings Of The 2018 World Wide Web Conference (WWW '18)*, 1491–1500.

Ravenscraft, E. (2020, December 18). Discord gaming parties are better than zoom meetings. *Wired*.

Recktenwald, D. (2017). Toward a transcription and analysis of live streaming on twitch. *Journal Of Pragmatics*, *115*, 68–81.

Reid, E. M. (1991). *Electropolis: Communication and Community on Internet Relay Chat* [University of Melbourne, Department of History].

Restream Team. (2021). *56 Streaming stats you should know in 2021*. Restream.

Rosenfeld, M. J., & Thomas, R. J. (2012). Searching for a mate: the rise of the internet as a social intermediary. *American Sociological Review*, *77*(4), 523–547.

Ruvic, D. (2021, November 9). Social media users could disable algorithms in new u.s. proposal. *Reuters*.

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, *55*(1), 68–78.

Sahil Verma, & Julia Rubin. (2018). Fairness definitions explained. *2018 IEEE/ACM International Workshop On Software Fairness (FairWare)*, 1–7.

Santora, J. (2022, March 15). *75 Live Streaming Stats Every Marketer Should Know in 2022*. Influencer Marketing Hub.

Schaffer, J., Giridhar, P., Jones, D., Höllerer, T., Abdelzaher, T., & O'donovan, J. (2015). Getting the message? a study of explanation interfaces for microblog data analysis. *Proceedings Of The 20th International Conference On Intelligent User Interfaces (IUI'15)*, 345–356.

Schirra, S., Sun, H., & Bentley, F. (2014). Together alone: motivations for live-tweeting a television series. *Proceedings Of The SIGCHI Conference On Human Factors In Computing Systems (CHI'14)*, 2441–2450.

Schütze, H., Manning, C. D., & Raghavan, P. (2008). *Introduction to Information Retrieval*. Cambridge University Press.

Scott, V. M., Mottarella, K. E., & Lavooy, M. (2007). Does virtual intimacy exist? a brief exploration into reported levels of intimacy in online relationships. *CyberPsychology & Behavior*, *9*(6), 759–761.

Seering, J., Flores, J. P., Savage, S., & Hammer, J. (2018). The social roles of bots: evaluating impact of bots on discussions in online communities. *Proceedings Of The ACM On Human-Computer Interaction*, *2*(CSCW), 1–29.

Seering, J., Kraut, R., & Dabbish, L. (2017). Shaping pro and anti-social behavior on twitch through moderation and example-setting. *Proceedings Of The 2017 ACM Conference On Computer Supported Cooperative Work And Social Computing (CSCW '17)*, 111–125.

Shachaf, P., & Hara, N. (2010). Beyond vandalism: wikipedia trolls: *Journal Of Information Science*, *36*(3), 357–370.

Sharma, A., & Yan, B. (2013). Pairwise learning in recommendation: experiments with community recommendation on linkedin. *Proceedings Of The 7th ACM Conference On Recommender Systems*, 193–200.

Sheng, J. T., & Kairam, S. R. (2020). From virtual strangers to irl friends: relationship development in livestreaming communities on twitch. *Proceedings Of The ACM On Human-Computer Interaction*, *4*(CSCW2), 1–34.

Shneiderman, B. (2020). Human-centered artificial intelligence: reliable, safe & trustworthy. *International Journal Of Human–Computer Interaction*, *36*(6), 495–504.

Sprecher, S. (1998). Insiders' perspectives on reasons for attraction to a close other. *Social Psychology Quarterly*, *61*(4), 287–300.

Sprecher, S. (2014). Effects of actual (manipulated) and perceived similarity on liking in get-acquainted interactions: the role of communication. *Communication Monographs*, *81*(1), 4–27.

Sprecher, S., Treger, S., Fisher, A., Hilaire, N., & Grzybowski, M. (2015). Associations between self-expansion and actual and perceived (dis)similarity and their joint effects on attraction in initial interactions. *Self And Identity*, *14*(4), 369–389.

Sprecher, S., Treger, S., Hilaire, N., Fisher, A., & Hatfield, E. (2013). You validate me, you like me, you're fun, you expand me: "i'm yours!" *Current Research In Social Psychology*, *21*(Article 3).

Sprecher, S., Wenzel, A., & Harvey, J. (2008). *Handbook of Relationship Initiation* (S. Sprecher, A. Wenzel, & J. Harvey (Eds.)). Psychology Press.

Stenberg, D. (2021). *History of IRC (Internet Relay Chat)*.

Strauss, J. P. (1993). *An examination of the relationships between actual and perceived personality similarity (between the rater and the ratee), liking, and performance ratings* [The University of Iowa].

Struzek, D., Dickel, M., Randall, D., & Müller, C. (2020). How live streaming church services promotes social participation in rural areas. *Interactions*, *27*(1), 64–69.

Sunnafrank, M. (1983). Attitude similarity and interpersonal attraction in communication processes: in pursuit of an ephemeral influence. *Communication Monographs*, *50*(4), 273–284.

Sunnafrank, M. (1984). A communication-based perspective on attitude similarity and interpersonal attraction in early acquaintance. *Communication Monographs*, *51*(4), 372–380.

Sunnafrank, M. (1985). Attitude similarity and interpersonal attraction during early communicative relationships: a research note on the generalizability of findings to opposite-sex relationships. *Western Journal Of Speech Communication*, *49*(1), 73–80.

Sunnafrank, M. (1986). Communicative influences on perceived similarity and attraction: an expansion of the interpersonal goals perspective. *Western Journal Of Speech Communication*, *50*(2), 158–170.

Sunnafrank, M. (1992). On debunking the attitude similarity myth. *Communication Monographs*, *59*(2), 164–179.

Sunnafrank, M., & Miller, G. R. (1981). The role of initial conversations in determining attraction to similar and dissimilar strangers. *Human Communication Research*, *8*(1), 16–25.

Sunstein, C. R. (2018). *#Republic: Divided Democracy in the Age of Social Media*. Princeton University Press.

Taber, L., Baltaxe-Admony, L. B., & Weatherwax, K. (2020). What makes a live stream companion? animation, beats, and parasocial relationships. *Interactions*, *27*(1), 52–57.

Tang, J. C., Venolia, G., & Inkpen, K. M. (2016). Meerkat and periscopee: i stream, you stream, apps stream for live streams. *Proceedings Of The 2016 CHI Conference On Human Factors In Computing Systems*, 4770–4780.

Terveen, L., & McDonald, D. W. (2005). Social matching: a framework and research agenda. *ACM Transactions On Computer-Human Interaction*, *12*(3), 401–434.

Tidwell, N. D., Eastwick, P. W., & Finkel, E. J. (2013). Perceived, not actual, similarity predicts initial attraction in a live romantic context: evidence from the speed-dating paradigm. *Personal Relationships*, *20*(2), 199–215.

Tixier, M., & Lewkowicz, M. (2016). “Counting on the group” reconciling online and offline social support among older informal caregivers. *In Proceedings Of The 2016 CHI Conference On Human Factors In Computing Systems*, 3545–3558.

Treger, S., & Masciale, J. N. (2018). Domains of similarity and attraction in three types of relationships. *Interpersona: An International Journal On Personal Relationships*, *12*(2), 254–266.

Uthus, D. C., & Aha, D. W. (2013). Multiparticipant chat analysis: a survey. *Artificial Intelligence*, *199–200*, 106–121.

Van, M., & Brynjolfsson, A. E. (1996). Electronic communities: global village or cyberbalkans? *Proceedings Of The 17th International Conference On Information Systems*, 32.

Viegas, F. B., & Donath, J. S. (1999). Chat circles. *Proceedings Of Conference On Human Factors In Computing Systems*, 9–16.

Vione, K. C. (2016). *The effect of actual and inferred value similarity on interpersonal liking* [Cardiff University].

Waldron, V. R., & Applegate, J. L. (1998). Similarity in the use of person-centered tactics: effects on social attraction and persuasiveness in dyadic verbal disagreements. *Communication Reports*, *21*(1), 155–165.

Walther, J. B. (1996). Computer-mediated communication: impersonal, interpersonal, and hyperpersonal interaction. *Communication Research*, *23*(1), 3–43.

Wang, D., Lee, Y.-C., & Fu, W.-T. (2019). “I love the feeling of being on stage, but i become greedy”: exploring the impact of monetary incentives on live streamers’ social interactions and streaming content. *Proceedings Of The ACM On Human-Computer Interaction*, *3*(CSCW), 1–24.

Wang, Z., Liao, J., Cao, Q., Qi, H., & Wang, Z. (2015). Friendbook: a semantic-based friend recommendation system for social networks. *IEEE Transactions On Mobile Computing*, *14*(3), 538–551.

Wee, J., & Lee, J. (2017). With whom do you feel most intimate?: exploring the quality of facebook friendships in relation to similarities and interaction behaviors. *PLOS ONE*, *12*(4), e0176319.

Weisz, J. D., Kiesler, S., Zhang, H., Ren, Y., Kraut, R. E., & Konstan, J. A. (2007). Watching together: integrating text chat with video. *Proceedings Of The SIGCHI Conference On Human Factors In Computing Systems (CHI '07)*, 877.

Werner, C., & Parmelee, P. (1979). Similarity of activity preferences among friends: those who play together stay together. *Social Psychology Quarterly*, *42*(1), 62.

Werry, C. C. (1996). Linguistic and interactional features of internet relay chat. In S. C. Herring (Ed.), *Computer-Mediated Communication: Linguistic, social, and cross-cultural perspectives* (pp. 47–64). John Benjamins.

Whittaker, S., Terveen, L., Hill, W., & Cherny, L. (2003). The dynamics of mass interaction. In *From Usnet to Cowebs* (pp. 79–91). Springer, London.

Wohn, D. Y. (2019). Volunteer moderators in twitch micro communities: how they get involved, the roles they play, and the emotional labor they experience. *Proceedings Of The 2019 CHI Conference On Human Factors In Computing Systems (CHI'19)*, 1–13.

Wohn, D. Y., Freeman, G., & McLaughlin, C. (2018). Explaining viewers' emotional, instrumental, and financial support provision for live streamers. *Proceedings Of The 2018 CHI Conference On Human Factors In Computing Systems*, 1–13.

Xu, B., Gao, G., Fussell, S. R., & Cosley, D. (2014). Improving machine translation by showing two outputs. *Proceedings Of The SIGCHI Conference On Human Factors In Computing Systems (CHI'14)*, 3743–3746.

Xu, D., & Tian, Y. (2015). A comprehensive survey of clustering algorithms. *Annals Of Data Science 2015 2:2*, *2*(2), 165–193.

YouTube. (2018). *Get started live streaming | YouTube Help*.

Zeng, X., Li, J., Wang, L., Wong, K.-F., Zeng, X., Wang, L., & Wong, K.-F. (2022). Modeling global and local interactions for online conversation recommendation. *ACM Transactions On Information Systems*, *40*(3), 1–33.

Zhang, A. X., & Cranshaw, J. (2018). Making sense of group chat through collaborative tagging and summarization. *Proceedings Of The ACM On Human-Computer Interaction*, *2*(CSCW), 1–27.

Zhang, W., Wang, J., & Feng, W. (2013). Combining latent factor model with location features for event-based group recommendation. *Proceedings Of The 19th ACM SIGKDD International Conference On Knowledge Discovery And Data Mining*.

Zheng, L., Xie, X., & Chen, L. (2018). Context-aware academic collaborator recommendation. *Proceedings Of The 24th ACM SIGKDD International Conference On Knowledge Discovery & Data Mining*, *10*, 1870–1879.

Ziegele, M., & Quiring, O. (2011). Privacy in social network sites. *Privacy Online*, 175–189.

Zytko, D., & Devreugd, L. (2019). Designing a social matching system to connect academic researchers with local community collaborators. *Proceedings Of The ACM On Human-Computer Interaction*, *3*(GROUP), 1–15.

APPENDIX A

Interview Protocol for the Study in Chapter 3

Welcome participant

Thank you for participating in this study. We will be asking you a set of questions about your experiences with live streaming platforms. We will be using the results to inform the design of future live streaming tools.

Informed consent

Before we begin, I will need to get your signature on this consent form. Please read it carefully and ask me any questions. Then, when you are ready, please sign the form. (distribute consent forms; get signature)

We will audio record the interview. All the audio files are for research analysis only and all the files are kept confidentially in our principal investigator's computer. No third party could get access to the data except our research group. We won't put your name, email, or any other personal information in any publication. But we will use your quotes, I mean your words and sentences from the interview as an evidence to support our research findings in the publications. If you are uncomfortable with the record, we will not use it. However, although every reasonable effort has been taken, confidentiality during the actual Internet communication procedures cannot be guaranteed.

As I said before, in today's interview we will go through some of your previous experience in live streaming. Live streaming here I mean the streamers use camera to shoot real life, audience can interact with viewers, such as traveling, outdoor activities, talk. But not include gaming, news, sport etc. Did I make myself clear?

Interview questions

Breaking the ice

First I'd like to ask you some questions about your experiences with live streaming.

How often do you watch live-streaming?

About how long ago did you start watching live-streams?

What kinds of live streams do you watch?

What do you mean by (the certain category)?

Why are you interested in these live streams?

(Probes for additional types of live streams)

Which platforms do you use to watch live streams?

(A certain platform)

On what devices do you watch live streams?

How often do you use this platform?

Why do you use this platform?

What kinds of live streams do you generally watch on this platform?

Do you ever give reactions (hearts etc.) to the streamer when watching on this platform?

If so, how often? Why? If not, why?

Do you ever chat with other audience members on this platform?

If so, how often? Why? If not, why?

What kinds of things do you talk about?

Do you ever chat with the streamer on this platform?

If so, how often? Why? If not, why?

What kinds of things do you talk about?

What features of this platform do you like?

Are there any features of this platform you dislike?

(next platform)

Repeat the above question block

What about (platform not yet mentioned)? Have you ever used that one?

Why watch live stream instead of pre-recorded video? If archived, which do you prefer to watch?

Generally what do you do when watching live streams?

A positive live streaming experience

Now I'd like to ask you some questions about a particular livestream. Think of a live stream you watched that you thought the most engaging, interesting and enjoyable.

What made you decide to watch this stream?

Where were you when watching this live stream? When?

What was the live stream about?

What streaming platform were you using?

Did you watch the entire stream? About how long was it/did you watch? Did you start from beginning?

How many audience were there?

Did you engage with the streamer during the livestream?

Did you provide reactions to the streamer, like hearts?

Did you chat with the streamer?

Did you engage with other audience members?

Did you chat with other audience?

How often? Why?

Did the other audience members engage with each other or the streamer during the livestream?

What were they doing? What were they talking about?

How often? Why?

Were you doing anything else while you were watching the livestream?

If so, what? Why?

Why were you doing this while watching live streaming?

What made it engaging, interesting and enjoyable?

Was there anything else about it that you found engaging, interesting or enjoyable?

A negative live streaming experience

Now I'd like to ask you some questions about negative live-streaming experiences. Have you ever had an unpleasant experience while watching live streams? If so, think of the most recent time this has occurred.

If participants cannot come up with any ideas. Give them some hints: have you ever encounter flaming? Or you cannot catch up with the chat?

Where were you when watching this live stream? When?

What made you decide to watch this stream?

What was the live stream about?

What streaming platform were you using?

Did you watch the entire stream? About how long was it/did you watch? Did you start from beginning?

How many audiences were there?

Did you engage with the streamer or other audience members during the livestream?

Did you provide reactions to the streamer, like hearts?

Did you chat with the streamer?

Did you encounter any unpleasantness chatting with the streamer?

If so, what happened?

Did you react in any way to this unpleasantness? If so, how?

Did others react in any way to this unpleasantness? If so, how?

Did you chat with other audience members?

Did you encounter any unpleasantness chatting with other audience members?

If so, what happened?

Did you react in any way to this unpleasantness? If so, how?

Did others react in any way to this unpleasantness? If so, how?

Did the other audience members engage with each other or the streamer during the livestream?

What were they doing?

How often?

Did you notice any unpleasantness in the audience discussion?

If so, what happened?

Did you react in any way to this unpleasantness? If so, how?

Did others react in any way to this unpleasantness? If so, how?

Wrap up

Do you have other things you want to tell me about your live-streaming experiences?

Thank you

APPENDIX B

Study Materials in Chapter 4

Chat Preprocessing

Annotation on Individual Chats

Timestamp	Username	Message	Tag
12:45:58 AM	clonedclone	pure carbs	grocery food
12:51:29 AM	Phil_Mun	that's a big market	grocery other
12:54:20 AM	iocane	love her food excitement	personal
12:52:53 AM	pedroz0rd	NotLikeThis	irrelevant

Chatter Data After Annotation

Username	grocery food	grocery other	personal	irrelevant
clonedclone	0.54	0.22	0.04	0.20
Phil_Mun	0.03	0.37	0.12	0.48
iocane	0.34	0.03	0.58	0.05

Chats Examples

Random Grouping

Timestamp	UserName	Content
12:45:54 AM	omegatreez	diet ruined sadKEK
12:45:58 AM	clonedclone	pure carbs
12:46:00 AM	BringoYaDingus	9000MG SALT
12:46:04 AM	Barrymccochner	THATS FOR HER
12:46:21 AM	TheNerdyPenguins	peepoFat good carbs
12:46:27 AM	douglarce02	Kingcurapane Pog
12:46:38 AM	handbanana9000	SHES STREAMING WITH ESFAND PogU
12:46:42 AM	Barrymccochner	DRY MARKETS
12:47:09 AM	watching_homes	SOME DAY, NEVER

12:47:15 AM	Barrymccochner	HES NEVER TRYING IT
12:47:16 AM	xiaolu87	its not 2x spicy its ez
12:47:19 AM	k0u_33	steak
12:47:23 AM	k0u_33	beef
12:47:35 AM	k0u_33	thin
12:48:15 AM	BringoYaDingus	M E A T
12:48:17 AM	Nerfy39	Pure fat Kreygasm

Similarity-based Grouping

Timestamp	**UserName**	**Content**
12:50:58 AM	malicesouls	Its murica. Everything has sugar in it KKonaW
12:51:57 AM	ToeMayToe8	kimchi so good
12:52:10 AM	malicesouls	I would say kimchi is sour
12:52:13 AM	ToeMayToe8	spicy and vinegar
12:52:15 AM	BongZxy	its fermented cabbage
12:52:24 AM	blackcloudbtw	Spicy fermented cabbage
12:52:38 AM	AngeloJulius	tastes like picked garlic with some chili peppers and a little fishy
12:53:04 AM	Fi3ndi5h	beer and wine are fermented
12:53:15 AM	ToeMayToe8	means good for your gut
12:53:21 AM	Blazingtree69	Less fermented better for Americans
12:53:33 AM	BongZxy	kimchi == sauerkraut, same fermented cabbage technique
12:54:36 AM	BongZxy	rice cake, so hella carbs
12:56:00 AM	malicesouls	All the carb peepoFat
12:56:38 AM	Fi3ndi5h	high in sodium
12:56:49 AM	musicmanx	Steak is bad to eat everyday

Post-session Survey

Please answer the following questions based on your experience in this session.

What is your survey code 1? (Please ask the experimenter, it differs in every session)

What is your survey code 2? (Please ask the experimenter, it differs in every session)

What is your SONA ID?

How many live streams have you watched in this study so far?

- ○ One
- ○ Two
- ○ Three

Please indicate the extent to which you agree with each of the following statement.
(1 - Very Low, 7 - Very High)

	1	2	3	4	5	6	7
I had fun watching the live stream	○	○	○	○	○	○	○
I enjoy watching the live stream	○	○	○	○	○	○	○
I would like to watch this streamer's live stream in the future.	○	○	○	○	○	○	○
I would like to watch similar types of live streams in the future	○	○	○	○	○	○	○

Experience of watching and chatting. Please indicate the extent to which you agree with each of the following statement.
(1 - Very Low, 7 - Very High)

	1	2	3	4	5	6	7
What level of success did you have in completing the task?	○	○	○	○	○	○	○
How much effort did you have to put into accomplishing the task?	○	○	○	○	○	○	○
How much thinking, deciding, calculating, remembering, looking, searching was required in this task?	○	○	○	○	○	○	○
How pressured for time did you feel when you were completing the exercise?	○	○	○	○	○	○	○
How insecure, discouraged, irritated, stressed and annoyed were you?	○	○	○	○	○	○	○

Please indicate the extent to which you agree with each of the following statement.
(1 - Strongly Disagree, 7 - Strongly Agree)

	1	2	3	4	5	6	7
I forgot about my immediate surroundings while viewing.	○	○	○	○	○	○	○
My watching experience was rewarding.	○	○	○	○	○	○	○
The sun rotates around the earth.	○	○	○	○	○	○	○
I felt interested in my watching task.	○	○	○	○	○	○	○
I was really drawn into my watching task.	○	○	○	○	○	○	○
Watching these videos was worthwhile.	○	○	○	○	○	○	○
I was so involved in the watching task that I lost track of time.	○	○	○	○	○	○	○
I felt involved in this watching task.	○	○	○	○	○	○	○
I wanted to continue watching out of curiosity.	○	○	○	○	○	○	○

Please indicate the extent to which you agree with each of the following statement.
(1 - Strongly Disagree, 7 - Strongly Agree)

	1	2	3	4	5	6	7
Overall, I am satisfied with this live chat tool.	○	○	○	○	○	○	○
I felt comfortable when watching this live stream with this live chat tool.	○	○	○	○	○	○	○
It was easy to interact with others using this live chat tool.	○	○	○	○	○	○	○

Please indicate the extent to which you agree with each of the following statement.
(1 - Strongly Disagree, 7 - Strongly Agree)

	1	2	3	4	5	6	7
I was able to read all the text chat messages.	○	○	○	○	○	○	○
I could reply to messages easily.	○	○	○	○	○	○	○
I felt the balance between too few and too many messages.	○	○	○	○	○	○	○

Live Chat - Please indicate the extent to which you agree with each of the following statement.
(1 - Strongly Disagree, 7 - Strongly Agree)

	1	2	3	4	5	6	7
I thought the conversation with others was **useful** for me.	○	○	○	○	○	○	○
I had so much **fun** talking with others when watching this live stream.	○	○	○	○	○	○	○
I **enjoyed** the converastion with others when watching this live stream.	○	○	○	○	○	○	○
I thought the conversation with other was **boring**.	○	○	○	○	○	○	○
I thought the conversation with others was **meaningful**.	○	○	○	○	○	○	○
I thought the conversation was **enjoyable**.	○	○	○	○	○	○	○
I thought the conversation with others was **a waste of time**.	○	○	○	○	○	○	○
The conversation did **NOT** hold my attention at all.	○	○	○	○	○	○	○

Please indicate the extent to which you agree with each of the following statement about **the people who watched the live stream with you just now**.
(1 - Strongly Disagree, 7 - Strongly Agree)

	1	2	3	4	5	6	7
I felt **close** with them when watching this live stream.	○	○	○	○	○	○	○
I would like to be friends with them.	○	○	○	○	○	○	○
I am interested in learing more about them.	○	○	○	○	○	○	○
I like these people.	○	○	○	○	○	○	○
I felt there was a strong sense of community in viewing this stream.	○	○	○	○	○	○	○
They have shared interests with me.	○	○	○	○	○	○	○
They are similar to me.	○	○	○	○	○	○	○
I would like to interact with them more in the future.	○	○	○	○	○	○	○

Final Survey

We are almost there! Please take a few minutes to complete the final survey. Thanks!

Generally, how often do you watch live streams?

- ○ Daily
- ○ 2-3 times a week
- ○ 4-6 times a week
- ○ Once a week
- ○ 2-3 times a month
- ○ Once a month
- ○ Never

Generally, when watching live streams, how often do you participate in the live chat?

- ○ Always
- ○ Most of the time
- ○ About half the time
- ○ Sometimes
- ○ Never
- ○ I do not watch live streams

Generally, I see the live chat as:

- ☐ A tool to comment on the streamer (do not expect response)
- ☐ A tool to talk to the streamer (expect interaction)
- ☐ A tool to communicate with other viewers
- ☐ A tool to express myself (including something not related to the live stream content)
- ☐ A tool to know how other viewers think about this live stream
- ☐ Other, please specify ______

Pleae briefly describe the three live streams you watched just now.

How much did you pay attention to the scenary in the travel live stream?

- A great deal
- A lot
- A moderate amount
- A little
- None at all

How much do you think viewers in the chat talk about the scenary in the travel live stream?

- A great deal
- A lot
- A moderate amount
- A little
- None at all

How much did you pay attention to the food in the grocery shopping live stream?

- A great deal
- A lot
- A moderate amount
- A little
- None at all

How much do you think viewers in the chat talk about the food in the grocery shopping live stream?

- A great deal
- A lot
- A moderate amount
- A little
- None at all

How much did you pay attention to the knowledge shared in the brain scientist live stream?

- ◯ A great deal
- ◯ A lot
- ◯ A moderate amount
- ◯ A little
- ◯ None at all

How much do you think viewers in the chat talk about the knowledge shared in the brain scientist live stream?

- ◯ A great deal
- ◯ A lot
- ◯ A moderate amount
- ◯ A little
- ◯ None at all

Here are a number of personality traits that may or may not apply to you.
Please indicate the extent to which you agree or disagree with each statement. You should rate the extent to which the pair of traits applies to you, even if one characteristic applies more strongly than the other.

I see myself as:

	Strongly Disagree	Disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Agree	Strongly agree
Extraverted, ethusiastic	○	○	○	○	○	○	○
Critical, quarrelsome	○	○	○	○	○	○	○
Dependable, self-disciplined	○	○	○	○	○	○	○
Anxious, easily upset	○	○	○	○	○	○	○
Open to new experiences, complex	○	○	○	○	○	○	○
Reserved, quiet	○	○	○	○	○	○	○
Sympathetic, warm	○	○	○	○	○	○	○
Disorganized, careless	○	○	○	○	○	○	○
Calm, emotionally stable	○	○	○	○	○	○	○
Conventional, uncreative	○	○	○	○	○	○	○

Please indicate to what extent you agree or disagree with the following statement.
(1 - Strongly Disagree, 5 - Strongly Agree)

	1 - Strongly Disagree	2 - Disagree	3 - Neither agree nor disagree	4 - Agree	5 - Strongly Agree
I will be able to achieve most of the goals that I have set for myself.	○	○	○	○	○
When facing difficult tasks, I am confident that I will accomplish them.	○	○	○	○	○
In general, I think I can obtain outcomes that are important to me.	○	○	○	○	○
I believe I can succeed at most any endeavor to which I set my mind.	○	○	○	○	○
I will be able to successfully overcome many challenges.	○	○	○	○	○
I am confident that I can perform effectively on many different tasks.	○	○	○	○	○
I can do most tasks very well.	○	○	○	○	○
Even when things are tough, I can perform quite well.	○	○	○	○	○

What year were you born?

What is your gender identity?

- Male
- Female
- Other

What is your ethnicity?

- White
- Black or African American
- American Indian or Alaska Native
- Asian
- Native Hawaiian or Pacific Islander
- Other

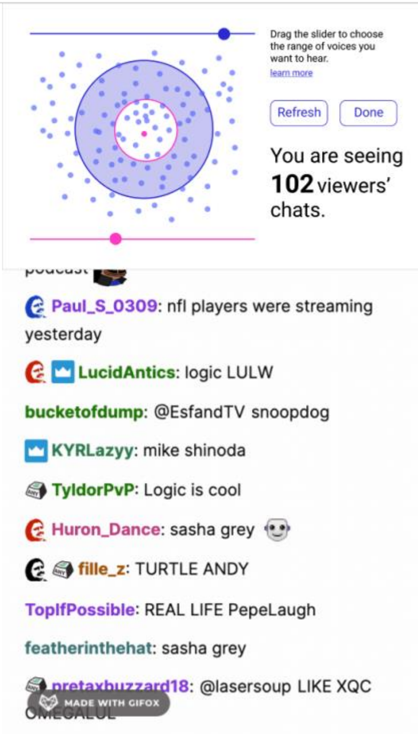
What is your SONA ID?

What is your username when you signed upt the chrome extension?

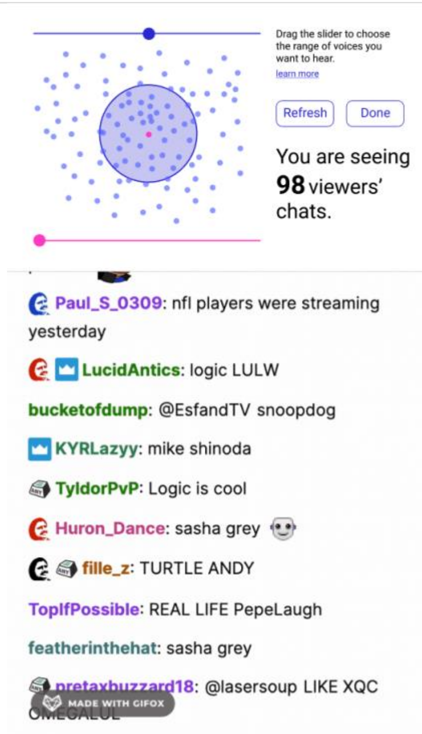
APPENDIX C

Design Iterations in Chapter 5

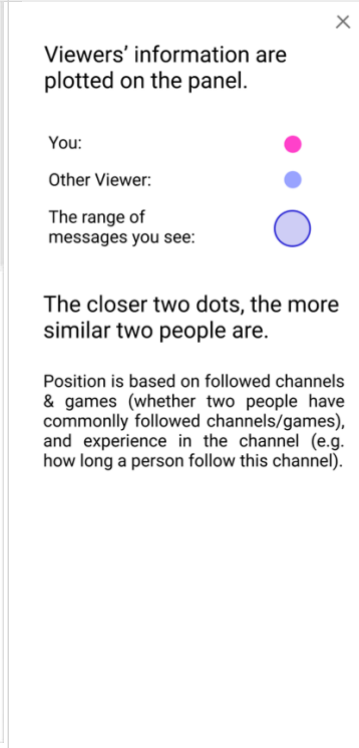
Iteration 3: Asymmetric Grouping, more interactions and explanations

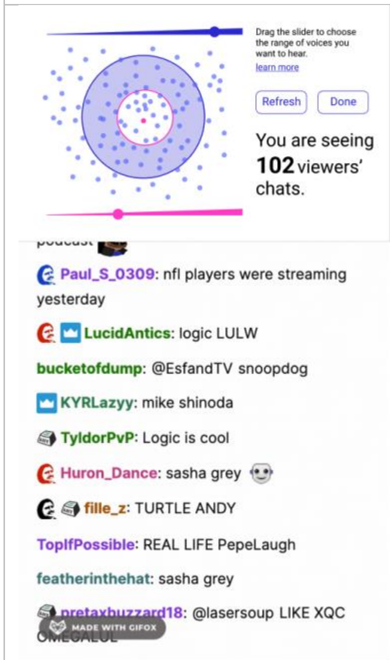
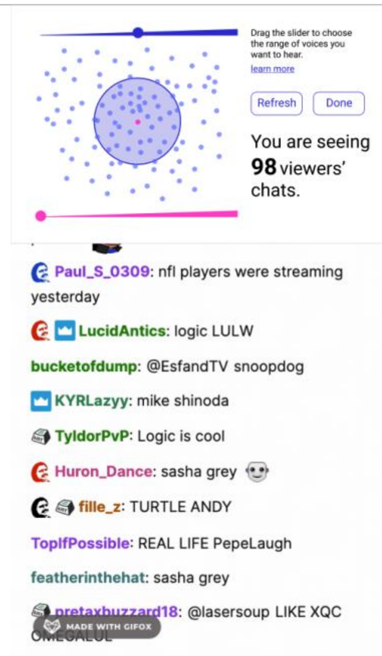
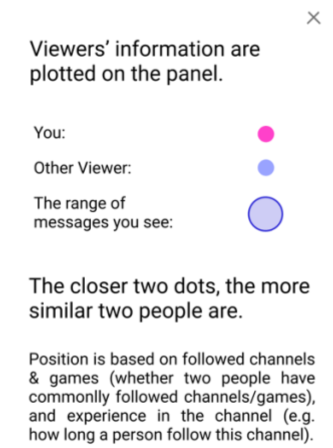


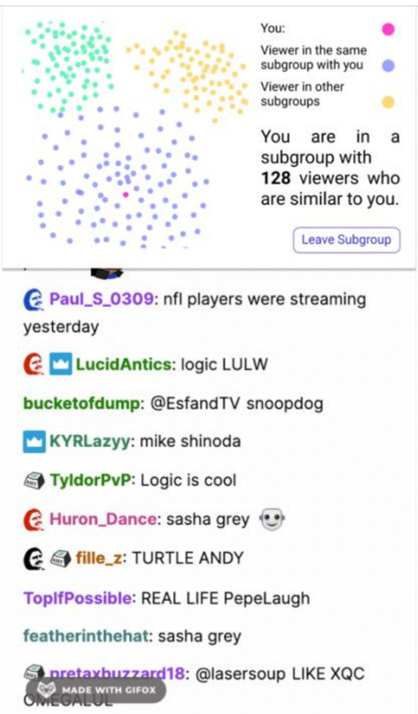
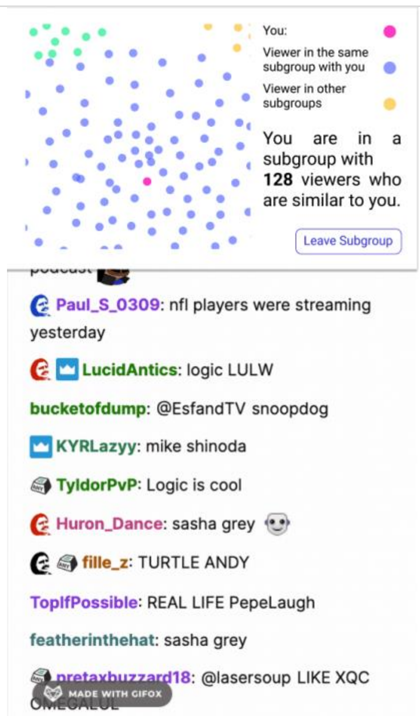
Iteration 3: Adjusted circle size



Iteration 3: Explanation after clicking "learn more"



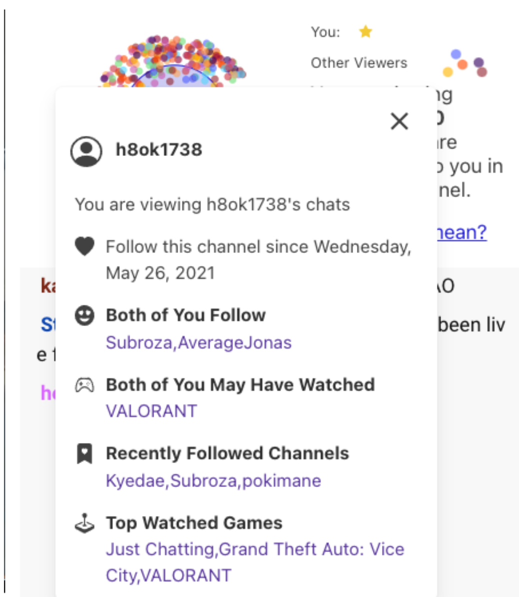
<table>
<tr><th>Iteration 4: Asymmetric Grouping, highlight slider directions</th><th>Iteration 4: Adjusted circle size</th><th>Iteration 4: Explanation after clicking "learn more"</th></tr>
<tr><td>
</td><td>
</td><td>
</td></tr>
</table>

<table>
<tr><th>Iteration 5: Symmetric Grouping</th><th>Iteration 5: Symmetric Grouping – Zoom in</th></tr>
<tr><td>
</td><td>
</td></tr>
</table>

APPENDIX D

Chatbuddies Interfaces

Hover on in-group viewer



Hover on out-group viewer

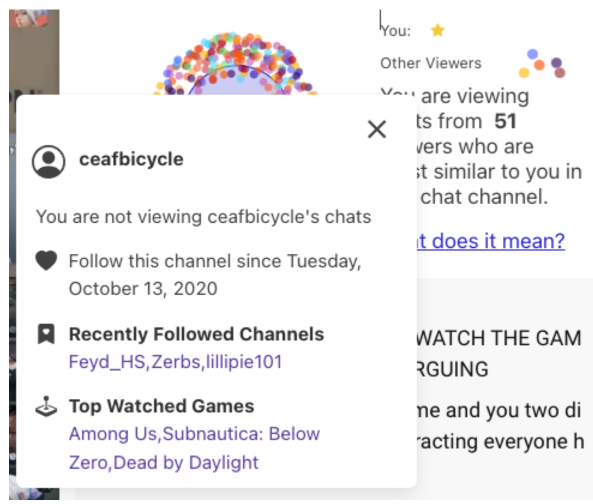
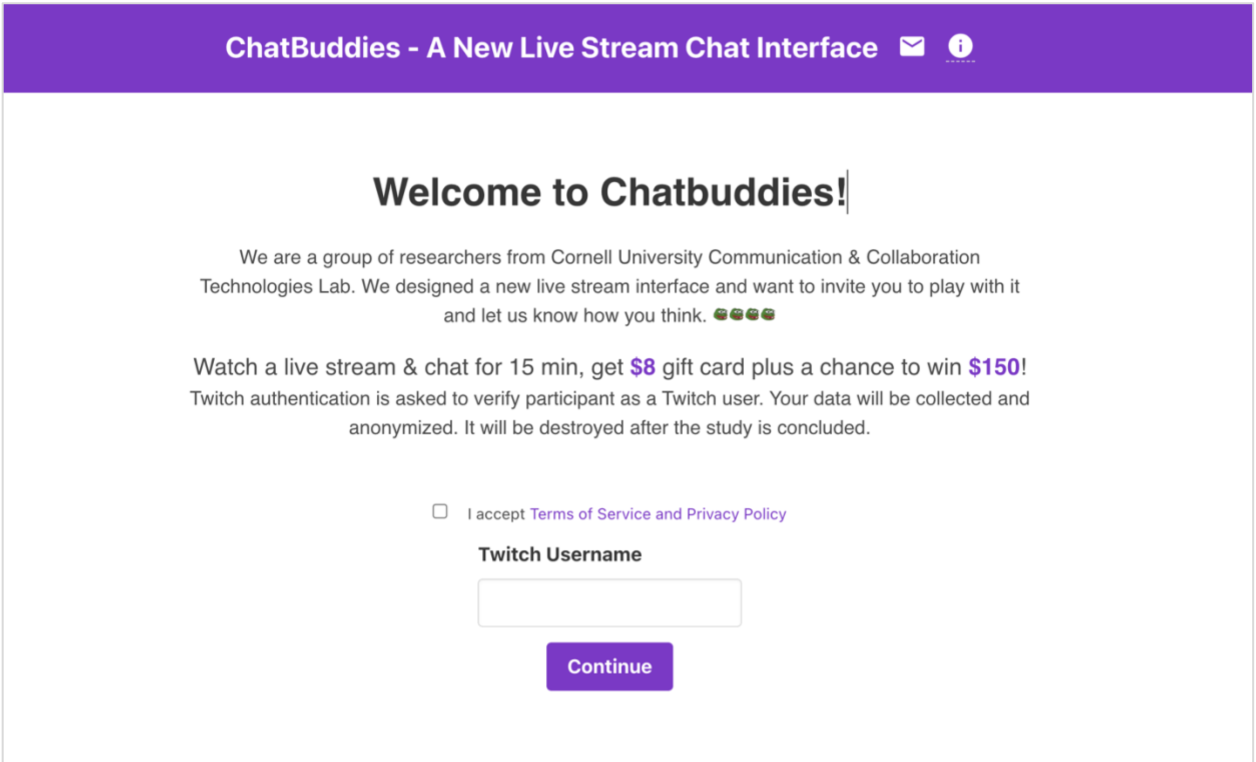


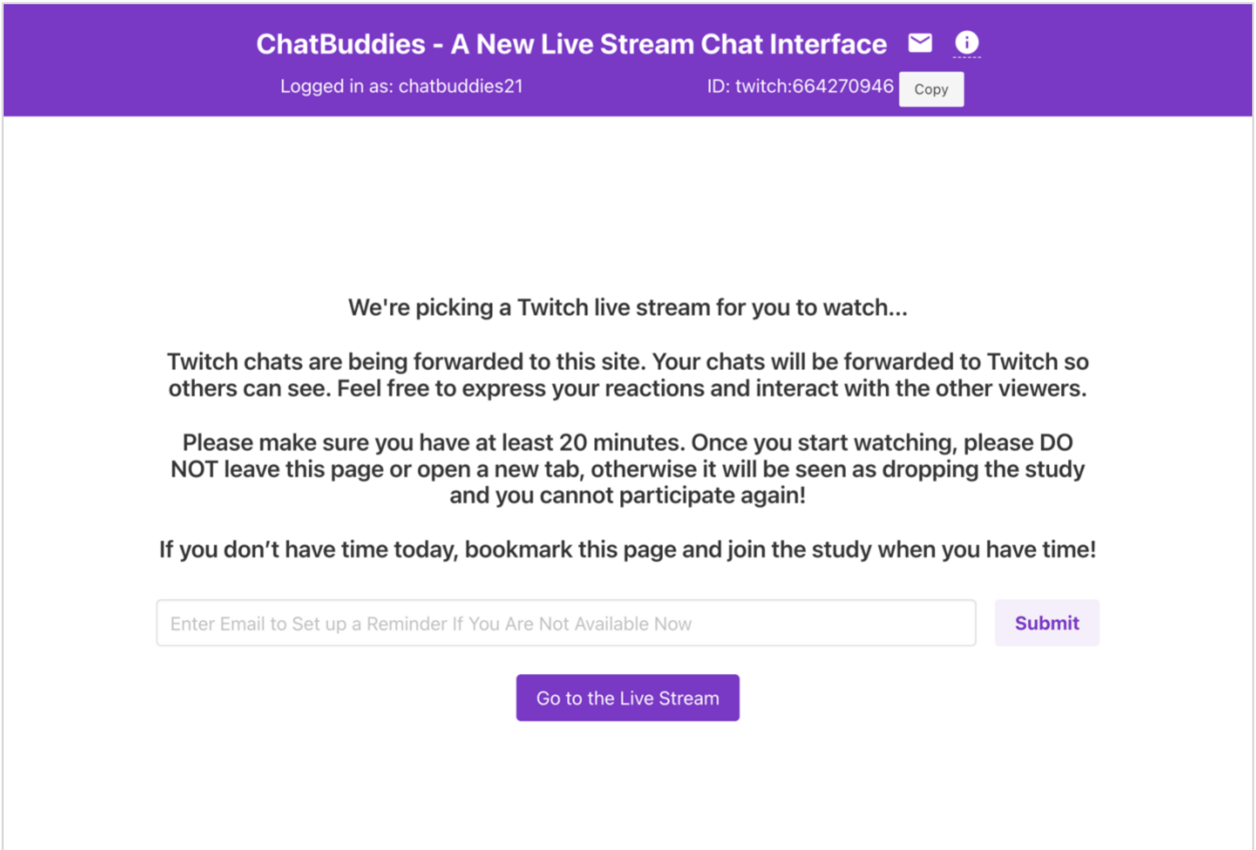
Figure 19: Hover on the scatter plot. In-group viewer (left) and out-group viewer (right)

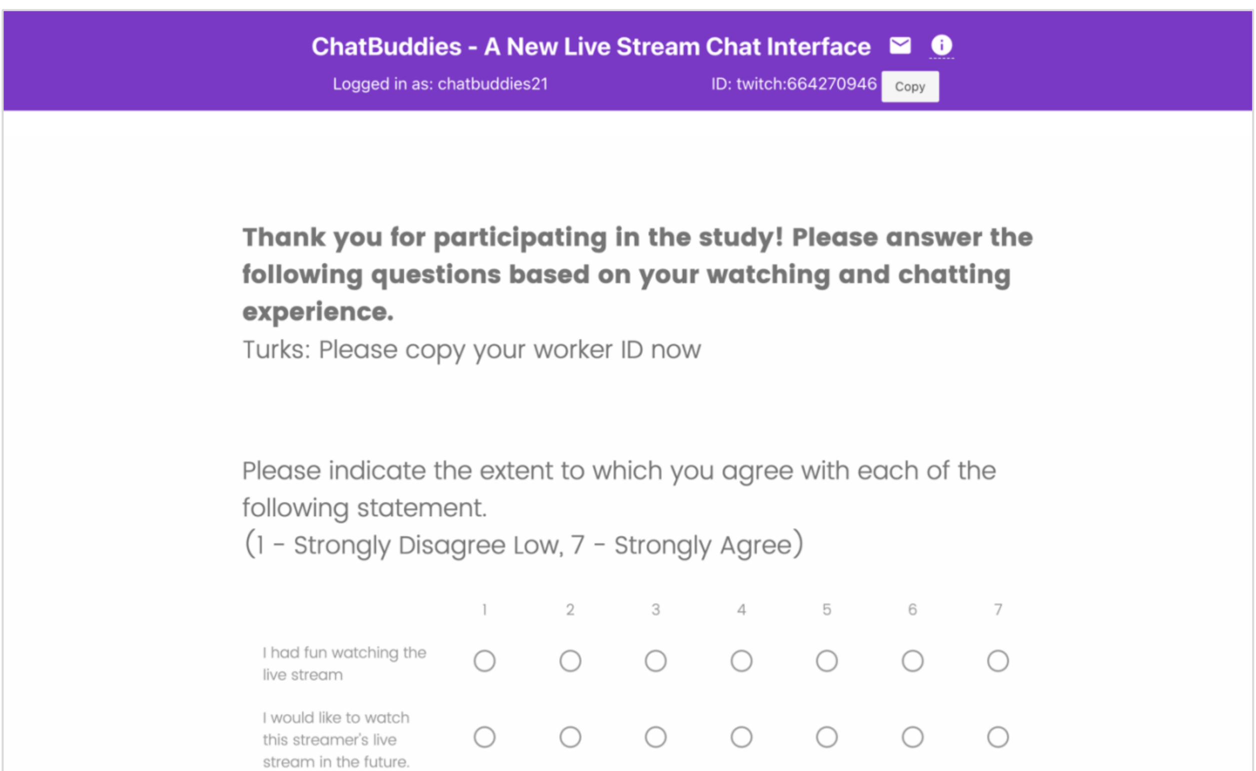
Question	Symmetric Grouping	Asymmetric Grouping
What does the color of dots mean?	Chatbuddies assign viewers to subgroups based on their similarity. Viewers in the same group have the same color. (Symmetric Similar) Chatbuddies assign viewers to subgroups randomly. Viewers in the same group have the same color. (Symmetric Random)	The color of dots is randomly assigned.
What does the position of dots mean?	In one subgroup, the closer two dots, the more similar these two individuals are. Positions of subgroups are randomly assigned (yellow group and green group are closer does NOT mean members are more similar).	The closer two dots, the more similar these two individuals are. (Asymmetric Similar) The position of dots is randomly generated. (Asymmetric Random)
How is similarity computed?	Similarity is computed based on mutually followed channels, games played by these channels, how long a user follow the current streamer, and subscription tier of this streamer.	

Table 8. Text in the Pop-up Tips

Chatbuddies Experiment Interface




ChatBuddies - A New Live Stream Chat Interface
Logged in as: chatbuddies21
ID: twitch:664270946
Copy
We're picking a Twitch live stream for you to watch...
Twitch chats are being forwarded to this site. Your chats will be forwarded to Twitch so others can see. Feel free to express your reactions and interact with the other viewers.
Please make sure you have at least 20 minutes. Once you start watching, please DO NOT leave this page or open a new tab, otherwise it will be seen as dropping the study and you cannot participate again!
If you don't have time today, bookmark this page and join the study when you have time!
Enter Email to Set up a Reminder If You Are Not Available Now
Submit
Go to the Live Stream


ChatBuddies - A New Live Stream Chat Interface
Logged in as: chatbuddies21
ID: twitch:664270946
Copy
Thank you for participating in the study! Please answer the following questions based on your watching and chatting experience.
Turks: Please copy your worker ID now
Please indicate the extent to which you agree with each of the following statement.
(1 - Strongly Disagree Low, 7 - Strongly Agree)
1
2
3
4
5
6
7
I had fun watching the live stream
I would like to watch this streamer's live stream in the future.

APPENDIX E

Online Survey for the Study in Chapter 5

Pre-study survey

Generally, how often do you watch gaming live streams on Twitch?

- ○ Daily
- ○ 2-3 times a week
- ○ 4-6 times a week
- ○ Once a week
- ○ 2-3 times a month
- ○ Once a month
- ○ Never

Generally, when watching gaming live streams on Twitch, how often do you participate in the live chat?

- ○ Always
- ○ Most of the time
- ○ About half the time
- ○ Sometimes
- ○ Never
- ○ I do not watch gaming live streams

Why do you watch gaming live streams on Twitch? (Please check all that appy)

- ☐ Enjoy an interest
- ☐ Learn new things, get information
- ☐ Build social connections
- ☐ Keep updated about the streamer and other viewers
- ☐ [ ] Other

What is the viewer count of the gaming live streams you usually watch?

- ○ 1,000 - 10,000
- ○ 100 - 1,000
- ○ > 10,000
- ○ < 100

Generally, I see live chats in gaming live streams as:

- ☐ A tool to comment on the streamer (do not expect response)
- ☐ A tool to talk to the streamer (expect interaction)
- ☐ A tool to communicate with other viewers
- ☐ A tool to express myself (including something not related to the live stream content)
- ☐ A tool to know how other viewers think about this live stream
- ☐ Other, please specify

What year were you born?

Post-session survey

Thank you for participating in the study! Please answer the following questions based on your watching and chatting experience.

Please indicate the extent to which you agree with each of the following statement.
(1 - Strongly Disagree Low, 7 - Strongly Agree)

	1	2	3	4	5	6	7
I would like to watch similar types of live streams in the future	○	○	○	○	○	○	○
I would like to watch this streamer's live stream in the future.	○	○	○	○	○	○	○
I had fun watching the live stream	○	○	○	○	○	○	○
I enjoy watching the live stream	○	○	○	○	○	○	○

How did you split attention on video and chats?

○ fully on video
○ mostly on video
○ half video half chat
○ mostly on chat
○ fully on chat

Please indicate the extent to which you agree with each of the following statement. (1 - Strongly Disagree, 7 - Strongly Agree)

	1	2	3	4	5	6	7
I was really drawn into my watching experience.	○	○	○	○	○	○	○
I was aware of most events happening in the game	○	○	○	○	○	○	○
I was really drawn into my chatting experience.	○	○	○	○	○	○	○
I was aware of most messages in the chat.	○	○	○	○	○	○	○

Please indicate the extent to which you agree with each of the following statement.
(1 - Strongly Disagree, 7 - Strongly Agree)

	1	2	3	4	5	6	7
I **enjoyed** the converastion with others when watching this live stream.	○	○	○	○	○	○	○
The conversation did **NOT** hold my attention at all.	○	○	○	○	○	○	○
I had so much **fun** talking with others when watching this live stream.	○	○	○	○	○	○	○
I thought the conversation was **enjoyable**.	○	○	○	○	○	○	○

Please indicate the extent to which you agree with each of the following statement about **the people who chatted with you (viewers in the circle) just now**.
(1 - Strongly Disagree, 7 - Strongly Agree)

	1	2	3	4	5	6	7
I felt **close** with them when watching this live stream.	○	○	○	○	○	○	○
I would like to be **friends** with them.	○	○	○	○	○	○	○
I am interested in **learning more** about them.	○	○	○	○	○	○	○
I **like** these people.	○	○	○	○	○	○	○

Please indicate the extent to which you agree with each of the following statement. (1 - Strongly Disagree, 7 - Strongly Agree)

	1	2	3	4	5	6	7
I felt there was a strong sense of community in viewing this stream.	○	○	○	○	○	○	○
I felt like a member of this live stream channel.	○	○	○	○	○	○	○

How much do you think you have in common with the other chatters (viewers in the circle) in terms of the following dimensions?. (1 - Not at all, 7 - A great deal)

	1	2	3	4	5	6	7
gaming preference	○	○	○	○	○	○	○
chat style	○	○	○	○	○	○	○
personality	○	○	○	○	○	○	○

Please tell us more about how you think of this live stream chat system.

Any Ideas for Improvements?

Is it the first time you participate in this study?

- ○ Yes
- ○ No

You saw chat messages from a subset of viewers. How did the system select the subset of viewers?

- ○ By random
- ○ By similarity
- ○ I don't think I saw chat messages from a subset of viewers

Final survey

How do you think of the two interfaces you have used in this study? Filter message by random vs by similarity.

Here are a number of personality traits that may or may not apply to you.
Please indicate the extent to which you agree or disagree with each statement. You should rate the extent to which the pair of traits applies to you, even if one characteristic applies more strongly than the other.

I see myself as:

	Strongly Disagree	Disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Agree	Strongly agree
Extraverted, ethusiastic	○	○	○	○	○	○	○
Critical, quarrelsome	○	○	○	○	○	○	○
Dependable, self-disciplined	○	○	○	○	○	○	○
Anxious, easily upset	○	○	○	○	○	○	○
Open to new experiences, complex	○	○	○	○	○	○	○
Reserved, quiet	○	○	○	○	○	○	○
Sympathetic, warm	○	○	○	○	○	○	○
Disorganized, careless	○	○	○	○	○	○	○
Calm, emotionally stable	○	○	○	○	○	○	○
Conventional, uncreative	○	○	○	○	○	○	○

What is your gender identity?

- ◯ Man
- ◯ Woman
- ◯ Non-binary
- ◯ Prefer not to disclose
- ◯ [ ] Prefer to self-descirbe

What is your ethnicity?

- ◯ Hispanic
- ◯ Non-hispanic

What is your race?

- ◯ White
- ◯ Black
- ◯ American Indian or Alaska Native
- ◯ Asian
- ◯ Native Hawaiian or Pacific Islander
- ◯ [ ] Other

Please tell us your email account to receive the Amazon gift card (Turkers please skip).

[ ]