Seeking Soulmate via Voice: Understanding Promises and Challenges of Online Synchronized Voice-Based Mobile Dating

Chenxinran Shen
University of British Columbia
Vancouver, Canada
elise.shen007@gmail.com

Yan Xu
University College Dublin
Dublin, Ireland
yen.hsu@outlook.com

Ray LC
City University of Hong Kong
Hong Kong, Hong Kong SAR
lc@raylc.org

Zhicong Lu*
City University of Hong Kong
Hong Kong, Hong Kong SAR
zhiconlu@cityu.edu.hk

ABSTRACT

Online dating has become a popular way for individuals to connect with potential romantic partners. Many dating apps use personal profiles that include a headshot and self-description, allowing users to present themselves and search for compatible matches. However, this traditional model often has limitations. In this study, we explore a non-traditional voice-based dating app called "Soul". Unlike traditional platforms that rely heavily on profile information, Soul facilitates user interactions through voice-based communication. We conducted semi-structured interviews with 18 dedicated Soul users to investigate how they engage with the platform and perceive themselves and others in this unique dating environment. Our findings indicate that the role of voice as a moderator influences impression management and shapes perceptions between the sender and the receiver of the voice. Additionally, the synchronous voice-based and community-based dating model offers benefits to users in the Chinese cultural context. Our study contributes to understanding the affordances introduced by voice-based interactions in online dating in China.

CCS CONCEPTS

• **Human-centered computing** → **Human computer interaction (HCI)**; *Empirical studies in HCI*.

KEYWORDS

Online dating, affordance, voice, social media, online communities

---

*Corresponding author.

1 INTRODUCTION

The widespread adoption of Mobile Dating Applications (MDAs) has coincided with the widespread use of smartphones. A study reveals that 38% of adults in the United States have tried MDAs at least once, highlighting the popularity of online dating as a method for people to meet potential romantic partners [49]. As more and more people join these platforms, the online dating industry has attracted substantial investments and has firmly established itself as a mature market [39, 50].

While HCI research has been exploring the affordances of online dating sites that heavily rely on text-based communication [27], voice-based online dating represents a relatively underexplored domain within the research landscape. Understanding the intricate mechanisms governing synchronous voice-based interactions in the context of online dating has received limited attention to date. By examining interactions mediated by voice, our study contributes to bridging this research gap, offering valuable insights that can benefit both researchers and dating platforms alike. These insights have the potential to drive improvements in the overall quality of communication within the realm of online dating [41]. We selected Soul as the primary voice-based dating platform for our study due to its popularity as one of the biggest voice-based dating platforms globally. As of January 2021, Soul had amassed 30 million monthly active users, ranking among the top apps in its category in China [5]. We conducted in-depth interviews with 18 dedicated Soul users and analyzed the data using thematic analysis [9].

In our study, we observed that Soul's emphasis on voice-based matchmaking has the potential to influence user behavior in several ways. Firstly, it directs user attention away from a primary focus on physical appearances, creating an environment that prioritizes communication. This shift appears to provide a more comfortable space for individuals who may have concerns about their physical appearance, potentially encouraging deeper conversations related to personalities, life experiences, and shared interests [46]. Additionally, voice communication allows users to convey responses and emotions through intonation and speaking style, which can enrich the quality of exchanges and add depth to conversations. Furthermore, the use of voice communication on Soul fosters a sense of naturalness and authenticity in interactions. With less reliance
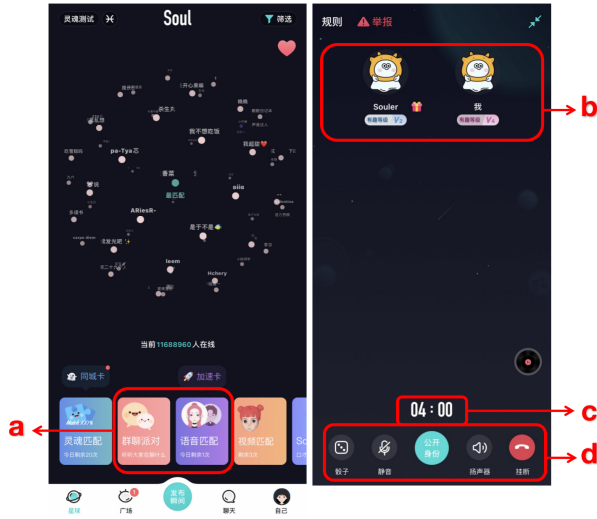
Figure 1: Left: The main user interface of Soul, which contains: (a) voice chat room and one-on-one voice matching. Right: The interface of voice matching, which contains: (b) Display of two users in an undisclosed state; (c) Indication of the remaining time for the chat; the chat will automatically end in four minutes if users choose not to proceed; (d) Various function buttons, including a dice for entertainment, the option to mute oneself, reveal one's profile, speaker controls, and the option to hang up the call.

on visual elements like photos and text, users may perceive these interactions as more genuine and personal, potentially leading to more positive online dating experiences.

Secondly, the inclusion of voice chat rooms within Soul plays a pivotal role in the quest for potential relationships. These rooms typically feature a host and multiple users participating in synchronous group voice calls. The presence of voice chat rooms diversifies the online dating landscape, offering users additional avenues to discover potential partners. These rooms come in various types, such as speed dating rooms, topic-based rooms, and general rooms, allowing users to join based on their interests and preferences. This further enriches the online dating experience on Soul. Inspired by Soul's popularity, we ask the following research questions:

- **RQ1**: What are the underlying motivations that drive users to choose voice-based online dating platforms?
- **RQ2**: What new affordances do voice-based interactions provide for self-presentation or self-disclosure separately in online dating?
- **RQ3**: How voice-based dating may benefit users under the cultural context of China?

Our research questions are relevant to the *hyperpersonal model* as they focus on the specific aspects of voice communication in online environments and its impact on relationship formation and social behavior. Our findings provide insights into the dynamics of voice-based dating within specific cultural practices. These include: 1) the influence of voice as a moderator in shaping perceptions and managing impressions between the sender and receiver in

communication; 2) the advantages that synchronous voice-based and community-based dating platforms provide to users within the Chinese cultural context. Our research contributes to the understanding of online dating communities that utilize voice as the primary modality, specifically within the unique cultural context of China, an East Asian country.

2 RELATED WORK

Our review centered on various aspects, including self-presentation, the design of online dating features, and the dynamics of dating communities. Additionally, we extended our investigation to encompass online dating cultures in non-Western countries, especially China, seeking a more profound comprehension of user behavior within the context of Soul.

2.1 Self-Presentation and Self-Disclosure in Online Dating Sites

In the early stages of self-presentation on online dating platforms, individuals aim to establish a favorable initial impression among other users. Impression management involves two key processes: first, the motivation to create a particular impression, and secondly, the actual construction of that impression [64]. There are various strategies to construct an ideal impression, such as crafting an appealing personal profile and engaging with others while adopting a curated persona [15, 64].

One approach to understanding deception in online dating is through signaling theory, which examines the connection between signals and the characteristics users wish to convey. Users express traits that may not be immediately evident and attempt to persuade others to perceive the credibility of these signals [12, 14, 53]. Previous research suggests that individuals who emit specific signals on online dating sites tend to be more appealing to other users [35]. Personal profiles serve as a consistent identity on the platform, aimed at attracting the attention of others [58]. On many online community platforms, creating a personal profile is essential for signaling [60]. To enhance their self-presentation and attract potential matches, users on online dating sites often provide more misleading information compared to other social media platforms [24, 39].

In traditional online dating apps, text-based communication is the primary mode through which potential partners interact. In essence, communication methods like text play a crucial role in facilitating interaction and self-presentation among potential partners [16]. However, voice can also be a significant factor in socializing with strangers, extending beyond dating applications. For instance, in emerging paid gaming teammate communities, multiplayer gamers often select their teammates based on their preferred voice and are willing to pay to play with them [47].

The hyperpersonal model [61] suggests that Computer-Mediated Communication (CMC) has the potential to facilitate more favorable social interactions compared to face-to-face encounters. This model has been applied within the context of online dating to elucidate why interactions in online settings may surpass those occurring in face-to-face scenarios [3, 24, 44]. The underlying reason, according to the model, is that message senders can effectively utilize the concept of self-presentation. Furthermore, various CMC technologies, such as instant messaging and online forums, can influence both

the type and quantity of information individuals choose to share with others. These choices can significantly impact how people form impressions of one another [23, 61].

The hyperpersonal model comprises four key components: receiver, sender, channel, and feedback, each contributing to enhanced social interaction. *Receiver:* In CMC, receivers interpret messages without the influence of physical appearance and vocal cues, which are typically central to forming impressions. Instead, they focus more on the content and style of the message, often leading to an extrapolation of meaning. They also tend to over-interpret relational messages, attributing more personal interest to online disclosures than in offline settings [29]. *Sender:* Senders in CMC engage in selective self-presentation. They have the advantage of carefully choosing message elements that best represent the image they wish to convey, a process more controlled than spontaneous speech. This also includes the ability to hide any undesirable physical traits or behaviors. *Channel:* The communication medium itself allows for further optimization of messages. *Feedback:* Idealized perceptions are formed and responded to with self-enhancing messages, which are further optimized through the advantages offered by the communication channel. This process leads to communicators being perceived as, and eventually becoming, more socially desirable.

The hyperpersonal model was applied to the context of online dating to elucidate why users often report greater satisfaction in the online dating environment compared to offline dating. This increased satisfaction can be attributed to the improvement in communication and the reduced reliance on physical cues to establish connections [3]. Online dating platforms offer individuals the opportunity to present an idealized version of themselves [24, 44]. Traditional online dating sites often center user profiles around physical appearance, while platforms like Tinder employ the "swipe right vs. left" feature, which enables users to form initial impressions of others based on their visual attributes before initiating further communication. This visual-first approach in Tinder exemplifies how online dating profiles are designed in many cases. Prior research has revealed that the inclusion of photographs in online dating profiles can reduce the idealized impressions that are typically generated by text-only information. This occurs because photographs introduce additional cues related to reality into the communication [62]. On the other hand, text-only information is often perceived as more trustworthy in terms of conveying an ideal impression [59]. Our study contributes nuanced insights to the Hyperpersonal Model, particularly regarding how users engage in selective self-presentation and the formation of idealized impressions in voice-based dating contexts.

2.2 Voice-based Social Interaction

Similar to visual information, vocal cues can convey characteristics about an individual. Voice is interpretable in terms of personality [67], and judgments about others' character can be based on their vocal traits [32]. Zhang et.al. conducted interviews with 15 active social media users to explore their preferences regarding the use of synthesized voices to represent their profiles. Their findings reveal that users desire control over how their voice delivers content, including factors like personality and emotion, as these

prosodic variations can significantly influence their online personas and impression management [70]. Research on audio media space is limited. Jung et al. conducted interviews with Clubhouse users, a voice-centered social media platform, to understand their motivations, social networking, and conversational habits. They discovered that voice-based interaction effective in forming social relationships [31]. Ackerman et al. investigated the suitability of audio for shared media systems, focusing on Thunderwire, a high-quality audio media space. They concluded that audio alone can create a functional media space and a valuable social environment. However, they noted that users needed to adapt to various audio-only and system-specific conditions [1].

Both vocal and physical appearances can create favorable impressions. This effect is more pronounced when individuals experience only one channel (either seeing physical appearance or hearing the voice) compared to multiple channels (simultaneously seeing physical appearance and hearing the voice) [71]. Compared to text-only communication, which relies on users' imagination, people tend to utilize additional physical factors, such as physical characteristics, when evaluating potential romantic partners [56]. Online dating apps that incorporate voice communication provide users with tangible physical attributes like reactions, voice tone, and speech patterns to aid in assessing whether a potential partner is a match. Sprecher conducted an experiment comparing text-only and voice-only communication in non-romantic relationships, which revealed that parties engaged in voice communication felt a greater sense of intimacy [55]. In essence, any form of communication with an audible component tends to outperform text-only communication [4].

Limited research in voice communication provides insight into the hyperpersonal model's intricacies. Duthler et. al. conducted a study comparing email and voicemail requests. They discovered that email requests tend to be more polite than voicemails. This difference arises because email allows senders to use more complex language and additional phrases, enhancing the message's politeness. On the other hand, voicemail's inability to be edited limits the sender's capacity to refine their message. The researchers claimed their findings consistent with the predictions of the hyperpersonal model. [13]. Another study focused on the gaming community's different reactions to male and female voices. The research revealed that female voices received three times more negative comments compared to male voices or instances where no voice was used. This suggests that voice communication in CMC can sometimes lead to increased hostility or negativity, especially towards women. This phenomenon, termed "hypernegative effects", highlights the impact of voice characteristics in online interactions [33].

2.3 Online Dating in Non-Western Cultures

While dating apps are widely accepted and popular for seeking relationships in Western countries [10], they often encounter a less favorable reputation in non-Western societies, where negative perceptions about these platforms are prevalent [8, 25]. This reputational disparity can lead to heightened privacy concerns and a greater tendency to conceal one's personal identity. Despite developers presenting these MDAs as tools for meeting new people and forming friendships, they are often viewed by many as primarily designed for casual encounters. MDA users, particularly women,

can encounter challenges related to resisting social stigma and dealing with unwanted sexual advances while using these applications. This divergence in dating and family cultures leads to distinct attitudes toward online dating sites between Western and Eastern cultures. In Eastern countries, the majority of individuals not only consider romantic compatibility but also take into account family backgrounds when connecting with potential partners [66]. Central to Chinese family culture is the concept of filial piety, where children are expected to date or marry someone approved by their parents. This can sometimes conflict with the use of MDAs, which may go against traditional family norms [2].

In non-Western countries, users of MDAs often grapple with heightened privacy concerns due to societal stigmatization and family cultural pressures [8]. To address these concerns and reduce uncertainty about potential romantic partners, users employ strategies to verify the credibility of their matches [21]. However, users of traditional MDAs frequently experience a loss of control over their personal information, inadvertently exposing it to unintended audiences [6]. This disclosure of personal information in profiles also increases the risk of identity theft or stalking [22, 54]. Consequently, many users are hesitant to reveal their MDA usage to individuals within their social networks [45]. Moreover, for MDAs that rely on image-based matching, users often find it challenging to maintain complete control over who can access their data, further intensifying privacy concerns.

Lots of eastern countries have a collectivistic culture, with China being an example. The collectivistic culture tend to value family, friends and the group over individual. In one study about the social networking site use, researchers found that personal characteristics were less effective in predicting social networking site (SNS) use in China than in the US. This may because in the collectivistic culture, the importance of the family, friends and one's groups may be partly responsible for Chinese participants' lesser use of SNSs [28]. The collectivist culture in China has undergone significant changes due to economic and social transformations, impacting dating and relationships among young Chinese adults [63]. There's been a shift from collectivism toward individualism, leading to a preference for romantic relationships over traditional, practical marital relationships. Such transformations motivates the needs in MDA [7]. While the mainstream of society has a low acceptance on MDAs, Chinese users have a higher privacy concern in MDA use [51]. In a study focusing on the motivations and risks for Chinese users of MDAs in meeting strangers, researchers discovered that concerns about self-exposure to friends, professional networks, and the community significantly influence their reluctance to meet people offline for casual sex. This highlights the impact of cultural norms and privacy concerns on dating app usage in China, particularly in relation to personal safety and social reputation.

2.4 Online Dating Community and Matchmaking

The matching algorithms in online dating apps can result in users being paired with various types of potential partners, and studies have revealed that users tend to select partners who share similarities with themselves, a phenomenon known as homophily [18]. This highlights the significance of designing effective matching

mechanisms in MDAs to steer users toward potential partners. Many MDAs are designed with a focus on 'speed dating,' aiming to expedite the matching process and reduce the time spent on communication [40]. However, this emphasis on speed and efficiency may discourage the development of genuine relationships and potentially foster a concentration on physical appearances. For traditional MDAs, the discovery mechanisms for users typically prioritize geographical location [72]. Many MDAs also include profile pages and private messaging interfaces as key features to help users screen potential partners [17, 72]. However, the common process of "discovering profiles, receiving match notifications, and interacting through messaging" often leads to misunderstandings and discrepancies between the personalities of users and how they present themselves in their profiles. Photos and text-only communication may not be sufficient for users to fully understand each other [40]. According to Lee and Bruckman, online dating is closely intertwined with real-life dating activities, often mirroring real-world dating dynamics [34]. In addition to one-on-one interactions, online dating can foster a sense of community where users share emotions and exchange information with each other [39, 65].

Traditional online dating apps typically rely on asynchronous and visual communication [20, 26]. Asynchronous communication implies that conversations between users may be delayed over time, which often happens in text-based communication. This delay can occur if one party does not respond immediately to a message or if one party is offline when a message is received. Visual communication refers to the fact that traditional online dating users have an initial visual impression of each other before they engage in communication. Successful matches are often based on photos, allowing both parties to form an initial impression of each other's appearance before initiating formal communication.

Soul distinguishes itself from traditional apps by not relying on asynchronous and visual communication. In Soul's voice-matching process, users remain anonymous and do not have access to each other's profiles, eliminating the influence of visual impressions. Communication on Soul takes place through voice, facilitating real-time interaction without the delays associated with messaging. Soul also introduces a unique feature, the voice chat room, where members can engage in group-based voice communication.

3 METHOD

To explore user behavior on Soul, we utilized a semi-structured interview approach with 18 participants. Our methodology focused on conducting in-depth discussions to gain insights into their experiences. Following the interviews, we employed thematic analysis to methodically examine and interpret the data collected. This approach allowed us to identify and understand key themes and patterns in user behavior within the context of the Soul app.

3.1 Participants

For our research, we enlisted the participation of 18 committed users of the Soul app. We recruited participants by initiating conversations in topicless chat rooms, entering as regular users to mingle with others. After exchanging greetings and introducing ourselves, we mentioned our ongoing study on the use of Soul and asked if anyone was interested in participating. For those who

expressed interest, we arranged to connect with them via WeChat. To ensure that our sample was representative, we specifically chose individuals who had engaged in several voice-matching experiences and had dedicated more than 5 hours to participating in voice chat rooms on the app. The demographic breakdown of our participants is detailed in Table 1.

The active user base of Soul is typically comprised of individuals under 35 years old [69]. To ensure that our participants fell within this age range, our primary focus was on individuals aged between 18 and 35 during the recruitment process. Our final sample consisted of 18 participants, with ages ranging from 21 to 26, including 10 female participants and 8 male participants. Out of 18 participants, 1 had experience hosting a voice chat room. All participants were informed that the interview process would be conducted anonymously. When selecting participants, we adhered to the following criteria: first, all participants had to be dedicated Soul users with over three months of experience, and second, they needed to have experiences with traditional dating apps as well.

3.2 Procedure

The interviews were meticulously conducted using Zoom, supplemented by WeChat for some participants, to provide additional insights during the analysis and coding stages. At the beginning of the interview, participants were thoroughly briefed through the reading of a consent form. This form detailed critical aspects such as the nature of the questions to be asked, the methodologies for data handling and storage, and the ethical considerations including participant rights. To proceed with the interview, verbal consent from each participant was mandatory.

The duration of each interview was approximately one hour. Interview questions were structured in four key areas: i) motivations behind the participants' use of the Soul application; ii) experiences and feedback regarding the app's voice matching functionality; iii) nature and quality of interactions within Soul's group chat rooms; and iv) comparative evaluations of Soul against other MDAs.

3.3 Data Collection and Analysis

The interviews for our study were conducted remotely via Zoom and were recorded in video format, generating approximately 20 hours of recorded material. Each interview session was captured as a video recording and later transcribed verbatim for analysis. One of the interviewers actively used the Soul app, providing an insider perspective during the research process.

In the analysis phase, three out of the four researchers employed thematic analysis [11] on the transcribed text. The team's roles were well-defined: one member was tasked with transcribing the video recordings into text format, while the other two researchers focused on coding and reviewing the transcribed data. Our thematic analysis combined both inductive and deductive coding strategies.

The analytical process encompassed several critical steps. Initially, each reviewer independently generated codes from their assigned section of the transcript, resulting in an average of about 50 unique codes per transcript. These initial codes included terms such

as "fear of being recognized by surroundings", "imagination (identity)", "imagination (personality)", and "stereotypes in voice/speaking style". Subsequently, we identified the relationships between the codes and grouped these codes deductively into broader categories, such as "privacy concerns", "receiver envisioning sender" and "perceptions of voice". We considered previous literature and theories, such as Hyperpersonal Model when developing the coding scheme. The final step of our analysis involved identifying relationships between these categories to develop a comprehensive set of themes. This process led to the formation of 30 distinct categories, which included both inductive insights and deductive groupings.

3.4 Ethical Considerations and Positionality Statement

This research project, including all related materials (e.g., recruitment plan, consent form, and interview scripts), was reviewed and approved by the Ethical Review Board prior to the commencement of the study to ensure compliance with ethical guidelines and research integrity. Before participating in the interview, each participant was provided with a detailed consent form to read and comprehend, ensuring informed consent.

The research team acknowledges the potential influence of our personal genders, cultural backgrounds, and social identities on the study, recognizing the possibility of introducing biases. Our team comprises four researchers: three identify as men and one as a woman. Ethnically, we are all of Asian descent. In terms of our relationship with the platform "Soul", our experiences vary: one team member is a dedicated Soul user, one has used Soul once, and two had not used Soul before this project. This diversity in our experiences with the platform is acknowledged as a factor in our research approach.

4 THE MECHANICS OF SOUL

Soul provides two avenues for users to discover potential partners: voice matching and group chat rooms (Figure 1). Voice matching entails users being matched with potential partners based on their personal information and preferences, followed by engaging in voice communication to establish a connection. Conversely, group chat rooms enable users to partake in group discussions on diverse topics, potentially leading to encounters with potential partners through these conversations.

Similar to traditional dating platforms, Soul requests users to furnish fundamental details about themselves during the registration process. This information typically includes gender, interests, and personality traits. These details serve as the basis for Soul's matchmaking services. When a one-to-one pairing is established, both users' profiles remain concealed from each other during the initial five minutes of the voice chat. During this time, their sole means of evaluation is through their voices. If both participants express mutual interest in extending their interaction, they have the option to unveil their profiles to each other, granting them unrestricted access to voice communication.

Soul offers users a choice of three distinct types of voice chat rooms: speed-dating rooms, topic chat rooms, and topicless rooms.

- Speed dating rooms: The purpose of these rooms is to assist users with particular preferences in discovering potential

Table 1: Summary of Interviewees

ID	Gender	Age	Occupation	ID	Gender	Age	Occupation
P1	M	24	Student	P10	F	21	Student
P2	M	23	Intern	P11	F	22	Unemployed
P3	F	22	Student	P12	F	22	Student
P4	M	24	Student	P13	F	23	Designer
P5	F	21	Student	P14	M	24	Student
P6	M	24	Intern	P15	F	23	Teacher
P7	F	24	Student	P16	M	24	Office Clerk
P8	F	24	Student	P17	M	20	Student
P9	F	26	Teacher	P18	M	24	Student

matches. Typically, these rooms are overseen by a *host*, a dedicated user responsible for creating the group chat, moderating the conversation, and steering the discussion toward a specific topic. The host listens to users' preferences and aids in identifying suitable matches based on those criteria.

- Topic chat rooms: Revolve around specific subjects, like hobbies or books, and participants in these rooms engage in discussions related to that particular topic.
- Casual chat rooms: Open-ended, allowing users to engage in discussions about a wide range of topics without any specific theme or constraint.

The voice chat rooms typically consist of four main components: the topic, the host(s), the speaker(s), and the audience(s):

- Topic: Only the topic chat rooms and speed dating rooms has a topic. A valid topic can be a hobby or a shared experience.
- Speaker(s): The participant in the voice room who is allowed to open their microphone and speak is typically the person looking for potential matches.
- Host(s): A experienced user who manages the voice-chat room and help participants get matched. The host decides the topic of the room and who can be the speaker.
- Audience(s): Everyone else in the voice-chat room.

These elements interact to shape the dynamics of the group chat room within the voice-based online dating app Soul, as illustrated in Figure 2.

In our interviews, participants discussed using self-created voice tags on the Soul platform to aid in finding romantic partners. These tags, developed by the users themselves rather than Soul's product team, are not used for automatic matching or filtering. Instead, users independently create and use these tags. To find someone with a desired voice type on Soul, users join voice chat rooms to listen to hosts or participants. Hosts may highlight a user's voice type to attract attention. Additionally, speed-dating rooms on the platform often welcome users with specific voice types.

The Soul community values the use of imagination in voice interactions, enhancing conversational engagement. With limited visual information, users rely on voice to understand each other's identities. To help form mental images, descriptive labels for voice timbres are used, such as "middle-aged" or "youthful" for male voices, and "authoritative" or "girlish" for female voices. These categorizations aid in comprehension and connection, shaping user perceptions and interactions on the platform.

5 FINDINGS

In the upcoming sections, we will delve into our findings and examine the factors motivating users to embrace voice-based dating platforms. We will also explore the distinctive dynamics of voice-based community interactions, and investigate the role of perception and regulation within these voice interactions.

5.1 Motivations Behind the Use of Soul

5.1.1 Multi-modal Information Needs. Similar to users of other dating apps, the primary motivation for Soul users is to satisfy their social interaction needs. Voice communication offers users the opportunity to glean more information about their potential matches than they might from a meticulously crafted profile page. Since users are unable to access each other's profiles until a few minutes into the voice chat, they depend on their voices to create an initial impression. Voice timbre, speaking style, and other attributes conveyed through voice communication collectively contribute to this initial impression. As articulated by P3,

> *"It's easier for me to evaluate one person through voice chat. I used to meet someone on Tinder who identify himself as a highly educated person. We have good conversations by texting to each other. Once we had a voice chat, I noticed that he wasn't the person he pretended to be immediately because of the way he spoke. Chatting through voice helps me better discover what the person really be like."* (P3)

Some participants expressed that they would invest more effort in sustaining a conversation if they found the other user's voice pleasant. For instance, P7 recounted her experiences using other dating apps in contrast to Soul, stating,

> *"I used to use Tinder. I would not spend the effort to keep the conversation via text messages and will not feel bad about pausing it. On Soul, I will take the initiative to continue the conversation if I like the person's voice and talking style."* (P7)

5.1.2 Voice Communication Enhances Privacy. In our study, six female participants underscored the negative reputation associated with traditional dating apps, noting that a significant number of individuals view them as disreputable. They articulated the belief that Soul offers a safeguard for their privacy by initially concealing their physical appearance. P5 desired to present her authentic self

on dating sites while avoiding recognition by individuals who knew her personally. However, a text-based dating platform that required personal images in profiles did not align with her requirements,

> *"If people around me know that I am using this kind of app, it may damage my reputation. If I put my image on this kind of app, there's a chance that I could be recommended to people who know me...You have to upload a personal image in some apps. I wouldn't upload a fake selfie since I don't want to engage with someone who uses a fake selfie either."* (P5)

Voice-based dating apps, in comparison to text-based dating apps lacking profile images, offer a higher degree of authenticity while simultaneously providing enhanced privacy protection. P2 and P4 expressed the similar opinion. As mentioned by P4,

> *"I want to reveal part of my real self without revealing my identity to people who may know me personally. Voice-based dating platforms can strike this balance."*

5.1.3 Emotional Enrichment Through Voice Communication Heightens Intimacy. Some participants suggest that they could receive and express a richer emotion signals via voice communication. Some participants felt it was easier to express their attitudes to potential partners via voice compared to text, *"I don't need to meticulously craft my messages to avoid misinterpretation,"* as P19 emphasized, *"I simply express my feelings naturally in the conversation. It make things easier."* Some users may get more emotional and empathetic to others by having voice calls instead of text messages on Soul. The emotions in one's voices can trigger one's desire to continue the conversation,

> *"Once I match with someone on the app who is using it due to loneliness and a desire to find someone to talk to. I would immediately pause the conversation if it's through text because I consider it a waste of time. However, when I hear their voice, I can sense their loneliness, and it reminds me of how I felt when I had just graduated from college."* (P15)

5.2 Perceptions and Moderation in Voice Interactions

In our study, we found that during voice exchanges, users often prioritize the *timbre* and *style* of the other person's voice over the actual content of the conversation initially. This tendency allows them to form an impression of the person's personality and attitude based solely on their voice. We adopted the framework of the Hyperpersonal Model, defining speakers as receivers and listeners as senders. Our research highlights how both senders and receivers perceive each other through voice, or use their own voice for self-presentation. This is based on the concept of forming a *mental image* of a person through their voice. Additionally, we explored the role of group voice chat room hosts in facilitating and moderating these voice interactions.

5.2.1 Receiver: Envisioning the Individual via Their Voice. The user assumes the role of an information receiver when they are in the position of a listener. As voice is the central mode of interaction on Soul, users heavily depend on it to shape their perceptions and

assumptions about each other's identities. The voice timbre, in particular, serves as a potent cue that enables users to visualize the appearance and identities of others based on voices and real-life experiences. The cue can be timbre, accent, speech mannerisms, and tone. For instance, if a male user's voice has a "teen boy's" timbre, P3 would envision him as a college student. P4 mentioned that he tends to imagine a girl who is taller than him whenever he hears a female speaker with an accent from the northern regions of China, "they usually glow taller than use from south", he said.

When users hear a voice type they like, they tend to idealize the speaker. P16 suggested that based on personal experience, it's easy to imagine a specific type of ideal partner just from their voice. *"I've realized I particularly like a certain type of girl. In the past, I paid extra attention to girls of this type in real life and noticed a similarity in their voices,"* he said. When asked about the mental image he forms, P16 became a bit shy,

> *"You know, I like girls who speak slowly and gently, sometimes a bit abstractly... Then, I picture them as well-educated ladies, likely majoring in arts or social sciences... I often imagine them in the same way."*

P14 shared his experience about how people find him attractive because of his voice, *"Many girls I met on Soul told me that because of my soft voice, they would imagine that I am a handsome gentleman, which makes them more willing to spend time chatting with me. "* Similarly, P13 noted that the lack of visible profile in one-to-one pairings fuels her imagination, *"Everything I know about him comes from his voice, which lets me imagine all the positive possibilities about him. Such imagination motivates me to continue the conversation and learn more about him."* She also compared voice-based dating to text-based dating apps and highlighted Soul's advantages in online dating, *"It's more intuitive, and the information is richer, allowing me to build up a mental image of this person in just a minute."*

When asked how their mental images would evolve through further communication, and whether an idealized image could negatively impact their perception if it contradicted the sender's real identity or personality, both P13 and P16 said the biggest impact of the voice type was encouraging them to invest more time and effort in getting to know the sender. P16 mentioned he is open to modifying his initial perceptions based on ongoing conversations, *"If her voice initially leaves a good impression, I find it worthwhile to invest more time and effort to understand her better and include more details to imagine her personality."* On the other hand, P13 noted that she tends to let go of any idealized mental images formed based on the sender's voice once deeper conversations begin, *"His voice is what motivates me to engage in further conversations with him. But I believe the best way to truly know someone is to avoid making assumptions from the start. "*

These voice types serve as a convenient tool for users to convey their preferences for specific voice timbres in potential partners. In scenarios like the speed-dating room, a user can simply reference a desired voice type, allowing other users to quickly understand the type of voice timbre the user seeks in a potential partner. This streamlined communication aids users in effectively connecting with each other on the app.

Additionally, beyond forming impressions of each other's identities and personalities through voice, receivers can also discern the

sender's attitude from how they speak. P3 shared her experience of evaluating male users' attitudes in the speed dating room, *"I can sense their emotions from the way they talk. It's hard to describe, but I can feel it. Some talk arrogantly, making me feel as if they're doing me a favor by speaking to me."* P7 commented on her ability to gauge attitudes in topic chat rooms, *"If I sense that the discussion is turning into a debate, I'll leave the room immediately. People's tone changes when the conversation is no longer peaceful."*

5.2.2 Sender: Managing Self-Presentation via Voice. Some users may deliberately modify their voices as a strategic attempt to appear more appealing to potential partners. This alteration might include using vocal techniques to adjust the timbre or pitch of their voices or adopting specific speaking styles they believe will enhance their attractiveness to others. As mentioned by P2. *"I will lower my voice and speak more slowly when talking to people on the soul. I want to leave the impression that I am a gentle person with a delicate voice. "* P7 shared that she intentionally makes her voice sound "cute" by speaking in a higher pitch and softer tone, and by giggling when she laughs. She noted that the impression she creates in voice chats contrasts with the first impression she typically gives in person,

> *"I'm taller and larger than most girls around me, and many people I've met in life don't see me as someone they need to protect. At some point, I started acting more like a tomboy and felt embarrassed about any feminine behaviors."* (P7)

P7 also discussed how voice-based dating platforms allow her to present herself in the way she prefers. *"I want to be treated like one of the delicate and soft girls, so I speak with a pitched voice. I can talk the way 'soft girls' talk."* Most importantly, she feels more comfortable presenting herself on voice-based dating apps without showing visual information. *"I would feel embarrassed and ashamed if people saw my photo, because my appearance doesn't fit the typical cute and feminine girl, but that's how I want to be treated."*

Some users experience a lower sense of self-deception when they modify their voice and speaking style to present themselves, compared to editing their selfies for dating profiles on other platforms. *"I can choose to speak in that manner anytime, but I can't always look like my edited selfie. Sometimes I feel like I'm lying to myself about how I really look,"* P7 explained. P8 shares a similar viewpoint, *"I've used Tinder before and posted an edited selfie. I didn't feel too bad about it because it seemed like everyone was doing it. However, it always feels better when people are attracted to the real you."*

However, there also exist users who use vocal changer software to decorate their voice, *"I used to talk with someone whose voice sounds unnatural in a voice chat room,"* said by P17, *"I think she was using a vocal changer software, but it only makes her voice sound weird. It also can be a male who uses software to pretend to be female and make fun. Anyway, I don't like it."*

5.2.3 Host—The "catalyst" for group chat room. The host plays a pivotal role in group chat rooms, facilitating interaction, guiding conversations, and helping users find compatible partners. They also regulate who can speak to maintain focus and efficiency. In topic chat rooms, hosts initiate conversations and set topics to foster connections. They moderate discussions to ensure relevance and respect within the community.

The first role of the host is to facilitate efficient dating in speed dating rooms, which typically have a large number of users looking for quick connections. The host acts as an organizer, screening suitable users and promoting efficient dating within the room. *"The host needs to drive the vibe and organize discourse in all aspects. "*(P9)

The second responsibility of the host involves cultivating a sense of community and encouraging the development of friendships in various group chat rooms. Users often frequent the same group chat rooms, sharing their experiences and getting to know each other over time. Typically, when a host creates a room, a consistent group of members joins, and they gradually become familiar with each other, forming a social circle. The host regularly sets up group chat rooms, which serve as the foundation for building long-term friendships. As mentioned by P9,

> *"There is a user who is the owner of the group chat room. By chance, I entered his group chat room because his voice timbre is very pleasant. He opened the room every day. At that time, the other two girls included me entered his group chat room on time every day...and we became friends."* (P9)

In topic chat rooms, the host also facilitates connections between users. These chat rooms typically have around 4-5 participants, depending on the host's settings when creating the room. Unlike speed dating rooms, chat rooms have a more relaxed atmosphere, and users may already be engaged in lively discussions. When a new user enters a topic chat room, the host often invites them to use their microphone and encourages them to join the ongoing conversation. This allows new members to participate in the topic discussion, providing an opportunity to connect with potential partners or friends who share the same interests.

> *"Basically, in order to seek politeness, the first reaction of everyone is not to interrupt rashly, but to listen for a while and then interrupt... If the host is in, he may invite users to speak if he finds a new partner coming in. Then at this time, users are not good to refuse, so that its retention rate will be higher. So actually the host I feel quite necessary."* (P1)

Furthermore, most hosts in Soul group chat rooms are expected to have a pleasant voice, as they play an active role in moderating and managing the room. A host with an appealing voice can attract more users to stay in the room and increases the likelihood of user engagement and participation. *"Many people will stay in the room because of the voice timbre of the host. The good voice timbre of the host can attract more people to enter the room."* (P12)

5.2.4 The Experience of Voice Interaction. Voice communication allows users to perceive the natural responses of the other party, enabling them to decide whether to continue the conversation based on these signals. Users can gauge the other person's interest and engagement in the conversation, which helps them determine whether to pursue further interactions. As participant P18 mentioned, *"Using voice to communicate makes me feel real because we can directly hear the other person's words, tone, and intonation."* When one person shows interest in the other, their tone of voice may convey a positive attitude. These emotions can be detected

through the voice, with changes in pitch, volume, and other vocal cues indicating the speaker's emotional state.

In text chat, users have the option to delay or pretend not to have seen messages they do not want to respond to immediately. However, there is no such delay in Soul's voice communication, and users must respond promptly. This immediacy makes it challenging to maintain long-term impersonation. Additionally, the synchronous nature of voice communication enables users to quickly reject someone who makes excessive demands. P15 mentioned that she would hang up when encountering users she did not like. *"If someone keeps mentioning that he wants to see my photos during the voice dating process, I will end the chat and leave immediately."* (P15)

5.3 Various Voice-Based Community Interactions and their Dynamics

In addition to one-on-one voice chat, the voice-based dating platform we examined also provides group-based voice chat environments. This allows us to explore how users engage within a well-established, voice-based dating community and gain insights into their perceptions of such interactions. There are three types of group chat rooms available on Soul: speed-dating rooms, topic chat rooms, and casual chat rooms.

5.3.1 Matchmaking with Complex Requirements with the Help of a Host in Speed-Dating Room. In Soul's speed-dating room, the goal is to help users with specific and diverse preferences quickly discover potential matches, all under the guidance of a host. Here's how the process works: Female users take turns to be the speaker by activating their microphones and describing the characteristics of their ideal partners. Male users who believe they meet these requirements can request to speak with the female user. The female user then chooses one male user, and they leave the room to initiate a private conversation. Throughout this process, the host moderates the conversation and determines who can use their microphones to address the group. Typically, the entire matching process is completed in less than two minutes.

Utilizing a speed-dating room can be more efficient in comparison to one-to-one matching on Soul. The advantage lies in the user's ability to engage with multiple potential partners simultaneously and assess them concurrently with minimal effort, e.g.,

> *"You can meet multiple people and talk to them simultaneously. Since it is a group conversation with the help of the host, you can get to know them quickly without spending effort on keeping the conversation going".* (P7)

In the matchmaking process in speed-dating room, the voice is a important factor. Users often appraise each other by their voice timbre and talking style. Since the time is limited, Therefore, voice timbre is still an important factor in speed dating room. Because the speed of making friends in a speed dating room is very short, and the speaking style and personality are difficult to be shown in a very short time, then users can only produce the first impression through the voice timbre, and often good voice timbre will gain more attention. As mentioned by P9, *"All the girls or boys in the speed-dating room will have requirements. Generally, ninety percent of them hope that the other party has a pleasant voice timbre."*

Figure 2: The interface of the Group chat room contains: (a) the topic of the Group chat room, if there is one; (b) the text chat area; (c) the top icon is the host user, and the rest users are divided into users with and without the mic. The users must apply to the host to open the mic and talk in the room. The discrimination method is to see whether there is a microphone sign on the right, and the number of people in the room is displayed at the bottom. (d) function buttons, which enable users to trigger an animation (for entertainment), send virtual gifts, and apply for permission to open up the microphone.

5.3.2 Finding Romance through In-Depth Conversations in Topic Chat Room. The topic chat room on Soul is distinguished by its specific and in-depth themes compared to a casual chat room. Conversations within a topic chat room revolve around a particular subject, allowing users to establish personal connections with others during the discussion. The room's host typically sets the topic, guiding the direction of the dialogue. Topics in these chat rooms can vary widely, encompassing shared experiences or discussions about movies, for instance. Among these rooms, the emotional room is the most popular, where users share their past relationships, provide support, and offer advice to one another. According to P11, *"The people who share their past relationship are looking for comfort. Some people want other users to critique their ex-boyfriend or girlfriend together with them. However, the ultimate goal of using the emotional room will always be looking for a new relationship."*

The voice chat room provides a structure and purpose for the conversation, which may be lost when the conversation moves to a different platform. It's common for online dating users to connect on third-party social platforms like Instagram or WeChat after getting to know each other on the dating app. However, in our study, we observed that some users prefer to limit their interactions solely to the topic chat rooms on Soul. When conversations lack a specific topic, they can sometimes become stagnant or come to a halt. For instance, P1 shared an experience where members in a topic chat room ceased chatting after attempting to continue their conversation on WeChat, *"After transferring the conversation to the WeChat group, some people started talking less, and eventually nobody talked anymore."* P12 participant suspected that it may because it's easier to initialize voice chat on Soul, *"It's common to have voice calls on a voice-based dating site like Soul. It can feel more intimate and even inappropriate to request a voice call when using other platforms."*

5.3.3 Connecting with Others through Casual Chat Room. In casual chat rooms on Soul, the host facilitates the interaction, and participants can join the discussion at any time by activating their microphone. These chat rooms lack a specific topic and allow users to engage in free-form conversations. The atmosphere in casual chat rooms is typically relaxed, and users are often evaluated based on their speaking abilities and sense of humor rather than their voice timbre. It's crucial for users to be mindful of the chat room's ambiance and purpose and adapt their communication style accordingly. *"In casual chat rooms, I think it is acceptable if the voice is not bad, and interesting words are more attractive. Compared with speed dating rooms, casual chat room people are relatively few. Chat time will be longer."* (P13)

5.4 Negative Experiences with Soul

Soul, like many other online dating apps, also grapples with a common issue—user fatigue from seeking new connections. Several participants revealed that they turned to online dating apps out of boredom, using them to combat loneliness and pass the time. However, this very convenience of always having new people to meet can lead to user fatigue, as individuals may become restless and develop unrealistic expectations that the next person they encounter will be even better. As P6 explained:

> *"It's a bit tired online dating because you will meet so many people. You feel like you're always going to meet someone better, and you are always picking and choosing, which can make you tired."* (P6)

5.4.1 Stereotypes on Voice. Soul, as an online dating app centered around voice communication, places a significant emphasis on the impression created by a user's voice. As we mentioned in Section 5.3, users form diverse impressions of others based on their voices. Individuals with melodious voices often have an advantage in making positive impressions, increasing their chances of meeting potential partners. However, it's crucial to acknowledge that not everyone possesses a pleasant or refined voice. Soul may not provide an equally favorable experience for users with average or less appealing voices. Some users with non-standard accents also faced discrimination on the platform. Participants raised concerns about accents, generally associating them with lower education

levels. As mentioned by P9, *"Because I was in a speed dating room. I often meet people who require to match with people that speaks standard Mandarin without accent."* P5 also holds the same point, *"For example, if the person does not speak standard Mandarin and then speaks with a strange tone, I will not have the desire to make friends with that person person."*

It's important to note that regional differences naturally result in various accents across China. What one person in the north might regard as non-standard Mandarin could be perceived quite differently by someone from the south. Therefore, such perceptions of "standard" versus "non-standard" Mandarin are often subjective and influenced by regional biases. The underlying cause of accent discrimination can be attributed to the varied environments in which people are raised, leading to different linguistic experiences and perceptions [52].

5.4.2 Conformity Phenomena and Voice. Some users become overly focused on voice timbre and may attempt to intentionally alter their own voices. Due to the popularity of certain timbres, like the bass voice, individuals might strive to sound more like these favored timbres. However, this deliberate imitation can have unintended consequences. One female participant mentioned encountering male users who intentionally deepened their voices to mimic the bass type. Participants shared that these attempts often came across as insincere and forced, leading to negative perceptions of those individuals.

> *"I think some boys speak a little affectation. But I don't know how to put it there are some people who start out just trying to attract people maybe or something. Because some girls, they just hang up when they don't like it. So some boys may talk deeper and more magnetic at the beginning to attract people. When we get to the end of the conversation, he relaxes. The voice may not be as good, or as magnetic."* (P5)

6 DISCUSSION

Our research aimed to investigate the impact of real-time voice-based MDAs in China . We explored the process of forming and nurturing potential romantic relationships through one-to-one voice interactions and group chats.

6.1 Affordances of Real-Time Voice-based Interactions in Online Dating

The Hyperpersonal Model proposes that individuals tend to be more satisfied with online dating because they disclose less information about their potential partner, which can lead to idealizing that person. Our research supports the hyperpersonal model, particularly in the context of voice interactions. In these scenarios, the roles of the receiver and sender can be seen as fluid, with speakers acting as receivers and listeners as senders.

6.1.1 The Role of Voice in CMC for Dating and the Influence of the Host. In previous research examining the online dating community, researchers identified a need for more "novelty" in online dating [39]. This refers to users of online dating forums expressing a desire for more enjoyable and diverse forms of socialization, as opposed to the uniformity often found in these platforms. We propose that

voice-based dating addresses this need for novelty by introducing a different form of communication, thereby diversifying the user experience in the online dating landscape. Based on the real-time nature of voice CMC, we suggest that voice-based dating may enable users to form a mental image of their potential partner more quickly than text-based dating, due to the rich information conveyed through the voice channel. Additionally, receivers who have a mental image of an ideal partner may quickly associate the sender with this ideal when they hear a voice that matches their model. Previous research indicates that voice communication is often perceived as more trustworthy than text because of its richer cues [30, 48]. However, in the context of online dating, senders might intentionally alter their voice and speaking style for self-presentation. Therefore, we hypothesize that receivers are more likely to accept the sender's crafted self-presentation on voice-based dating sites than on text-based dating sites.

The presence of voice types on platforms like Soul plays a significant role in fostering the idealization of voices. Our investigation into Soul's platform reveals that the most commonly used labels for male voices are "teen-boy" and "bass," while female voices are frequently labeled as "youthful" and "authoritative lady." These labels describe the pitch of the voice, with "teen-boy" and "youthful" indicating a higher pitch and "bass" and "authoritative lady" denoting a lower pitch. This labeling system ensures that almost every user can find a label that resonates with their voice. Such categorization of voices into distinct labels influences user perceptions and expectations. Users tend to associate certain personality traits and characteristics with these voice types. For instance, a "teen-boy" or "youthful" label might evoke an image of someone young and energetic, while "bass" and "authoritative lady" could suggest a sense of maturity and confidence. As a result, these labels not only help users identify voices that they find appealing but also contribute to the idealization of the voice's owner based on the associated characteristics of the label, further exemplifying the impact of voice perception in online communication.

Hosts in voice-based dating platforms offer a more authentic and efficient matchmaking process, overcoming the limitations of traditional MDAs that focus on searchable attributes, and addressing the inefficiencies found in game-based dating environments. The presence of a host in voice-based dating facilitates a more genuine form of matchmaking. Traditional MDAs often compel users to select partners based on searchable attributes like income or religion [19], but daters prefer to assess experiential qualities like humor or rapport. Previous research has shown that game-based dating offers a more immersive experience than typical online forums [68], but it lacks efficiency and consistency. In the voice-based dating site we studied, the host efficiently manages matchmaking in speed dating rooms with complex requirements and guides meaningful discussions in topic chat rooms. They also help forge user connections and foster a sense of community. However, regulating voice communication in these group environments poses more challenges compared to other communication spaces.

6.1.2 Rationale behind Sender and Receiver Dynamics. The **receiver**, or the listener in this case, forms first impressions of the sender's (speaker's) identity and personality based on their voice. This perception is influenced by various vocal characteristics such as timbre,

accent, speech mannerisms, and tone, as detailed in Section 5.3.1 of our study. If the receiver finds the voice timbre appealing, they are inclined to idealize the sender's identity and personalities. This could be because they subconsciously link the voice to that of an ideal partner. This association significantly shapes their perception and expectations during the initial stages of interaction, demonstrating the impact of voice characteristics on the formation and development of interpersonal connections in voice-based CMC environments. The process of interpreting these vocal cues is largely subjective, depending heavily on the receiver's personal experiences and preferences. For instance, as outlined in Section 5.4.1, a receiver might choose to discontinue communication with a sender if they perceive the sender's speech as deviating from what they consider standard Mandarin. This decision is based on the receiver's interpretation and preference, demonstrating how personal biases and expectations in voice play a significant role in voice-based CMC.

In the context of voice-based online dating, **senders** often adjust their voice and speaking style to present an idealized version of themselves, aiming to be more attractive to potential partners. This modification is a conscious effort by senders to align their vocal presentation with how they imagine their ideal self would speak in a dating scenario. For example, a user like P7, who aspires to embody the persona of a "soft girl," might find it challenging to express this identity in real-life interactions. However, in the online space, they have the opportunity to alter their voice to fit this desired persona. By softening their tone or adopting speech patterns that they associate with the "soft girl" image, they present themselves in a way that they believe will be more appealing to others.

In the context of voice-based dating platforms, we suggest that senders might experience a lower sense of self-deception when presenting their ideal selves via voice compared to other types of self-presentation on dating sites. This could be attributed to the nature of voice communication, which, while allowing for some degree of modulation or alteration, still retains a core element of the individual's authentic self — their actual voice. Unlike text-based or visually dominated platforms where users can heavily curate or manipulate their representation (through edited texts, photos, or entirely fabricated profiles), voice-based platforms inherently require a more genuine aspect of the individual's identity. Even when modulating tone or adopting different speech patterns to align with an idealized persona, the underlying uniqueness of one's voice remains a more authentic trait that cannot be completely disguised or altered. In addition, a user can easily alter their voice and speaking style to fit in their idealized self-presentation at any time. This relative authenticity in voice communication may lead to a feeling of less self-deception for the sender. They might perceive their presentation as a more honest or genuine version of themselves, even if it is an idealized one.

6.2 Voice-based MDA in China

Individualism and collectivism cultures can influence user perceptions and behaviors towards online dating [42]. In a country characterized by a collectivistic culture, the nature of such cultures tends to favor shared activities, such as livestreaming, over solitary ones [36–38]. This preference may explain why community-based

dating could be particularly appealing to users of Chinese dating sites. Voice chat rooms offer a unique dynamic, allowing users the choice to either distinguish themselves through speaking up or to assimilate seamlessly with the group. This flexibility can foster a sense of unity and belonging, potentially making users feel more at ease. Furthermore, topic chat rooms on platforms like Soul can significantly contribute to the development of romantic relationships by providing a space where users can connect over shared interests and commonalities.

Chinese users may also have a higher privacy concerns in using MDAs. This is because MDAs have a lower acceptance to public in China compared to western countries [51]. For the MDAs that necessitate personal profile creation and image uploads pose privacy risks, they potentially leading to judgment and scrutiny from individuals in their real-life circles. Chinese users, especially women, may encounter increased risks when utilizing these platforms due to the conservative social attitudes surrounding dating, relationships, and marriage in Chinese culture [43, 57]. Research and previous findings suggest that some Chinese MDA users may worry about the potential adverse consequences of disclosing their use of MDAs for romantic purposes to their family, professional networks, and community. Even after marriage, some couples formed through MDAs might opt to conceal the origin of their relationship to avoid gossip and potential social stigma. These concerns underscore the challenges and hazards inherent in the intersection of modern technology and deeply entrenched cultural norms in China [8]. By opting for voice-based dating, users can significantly mitigate the risk of being recognized through personal images, potentially enhancing their sense of security and comfort on MDAs.

6.3 Implications for Design and HCI Research

We propose several design recommendations inspired by the voice-based dating platform, which could enhance existing dating sites that rely on textual and visual information. These suggestions also offer potential avenues for further research on online dating.

6.3.1 Use voice profile to enrich personal information.
To accommodate users who are hesitant to reveal their personal images or other sensitive information, Soul promotes the use of voice profiles as an alternative method for connecting with other users. A voice profile typically consists of a brief voice recording that serves as a self-introduction. This voice profile can be incorporated into the "swipe left, swipe right" matching mode commonly used in many MDAs to facilitate connections between users. By listening to a voice recording, users can gain insights into another person's personality and interests, allowing them to decide whether they want to establish a connection. Integrating voice recordings into this context serves several purposes. It enhances the authenticity and genuineness of a user's profile, eliminates potential discomfort or reluctance associated with sharing sensitive personal data, and provides an alternative means for users to connect and acquaint themselves with others while protecting their privacy.

6.3.2 Encourage diverse self-presentation types.
Online dating platforms should provide users with features that allow them to present themselves based on their dating needs and purposes, either authentically or ideally, and showcase different aspects of their personality and interests beyond their voice types. This might be achieved by incorporating multimedia content: Allowing users to include photos, videos, or audio recordings that showcase their interests and personality can provide a more holistic representation of themselves and encourage diverse self-presentation.

6.3.3 Counteract homogeneous standards.
Our research indicates that certain voice types are perceived as more attractive on voice-based dating platforms, leading users to alter their voices to fit these preferences. Online dating platforms should implement features that challenge users' homogeneous standards of voice by diversifying voice types of the recommended potential dates. This might be achieved through carefully designed algorithmic interventions. For example, platforms can develop algorithms that deliberately present users with voice types that defy stereotypical norms formed within the communities. By exposing users to a wider range of self-presentation types of voice, online dating platforms can encourage users to broaden their preferences and challenge their unconscious biases. The platforms should also consider highlighting diverse profiles with different voice types. For example, platforms can feature diverse profiles in promotional sections or recommendations to promote the inclusivity of some rare voice types. Platforms can potentially disrupt homogeneous standards and encourage user acceptance by showcasing users who do not adhere to traditional voice standards or self-presentation types.

6.3.4 Add voice chat room feature.
As discussed in section 4, even text-based MDAs that incorporate voice chat features may still encounter user reluctance. To address this challenge and promote increased participation in voice-based communication, one potential solution is to introduce dedicated voice chat rooms within text-based MDAs. The inclusion of dedicated voice chat rooms provides a designated space for voice interactions, which can effectively reduce users' psychological barriers and enhance their comfort and confidence in utilizing voice communication. This is particularly advantageous for individuals who may experience anxiety or self-consciousness when engaging in text-based chat or who simply prefer a more intimate and immersive mode of communication.

6.3.5 Date in a Group Setting .
Our research indicates that group dating settings, similar to those found in interest groups or discussion rooms on other dating sites, may offer distinct advantages to users. These platforms provide a valuable space for users to engage and interact with others who share similar interests. Furthermore, these communities might shift the focus away from visual aspects, encouraging users to concentrate more on the shared interests and the sense of community itself.

6.3.6 Considering Other Modalities Beyond Visual.
We also advocate for expanding research into dating sites that emphasize modalities beyond the visual aspect. While this project concentrates on a voice-based platform within our demographic confines in China, there is potential novelty in designing a new platform focused on voice or other modalities, which could make significant design contributions to the field.

6.4 Limitations

Our participant recruitment was limited to a non-Western context, as a mature voice-based dating platform was only available in China at the time of our study. All of our participants are young adults in their early to mid-20s. Therefore, how users from different age groups perceive voice-based dating experiences remains unknown and we leave it for future research.

7 CONCLUSION

We conducted an interview study to explore the motivations behind using voice-based online dating apps and to understand how users interact throughout the dating process. Our findings highlight that the incorporation of voice communication within the Soul platform fosters a deeper mutual understanding among potential romantic partners, as it allows users to readily discern each other's reactions and emotions. Beyond assisting users in finding romantic partners, the diverse dating processes and themed voice chat rooms on Soul also offer various avenues for users to forge new friendships. In summary, our research underscores the effectiveness and significance of incorporating voice communication into online dating, providing individuals with a meaningful means of connecting and getting acquainted. We anticipate that future research will further delve into the potential of novel communication methods within online dating applications.

REFERENCES

[1] Mark S. Ackerman, Brian Starr, Debby Hindus, and Scott D. Mainwaring. 1997. Hanging on the 'wire: A Field Study of an Audio-Only Media Space. *ACM Trans. Comput.-Hum. Interact.* 4, 1 (mar 1997), 39–66. https://doi.org/10.1145/244754.244756

[2] Adel Al-Dawood, Norah Abokhodair, Houda El Mimouni, and Svetlana Yarosh. 2017. "Against Marrying a Stranger" Marital Matchmaking Technologies in Saudi Arabia. In *Proceedings of the 2017 Conference on Designing Interactive Systems.* 1013–1024.

[3] Traci L Anderson and Tara M Emmers-Sommer. 2006. Predictors of relationship satisfaction in online romantic relationships. *Communication Studies* 57, 2 (2006), 153–172.

[4] Marjolijn L Antheunis, Alexander P Schouten, and Joseph B Walther. 2020. The hyperpersonal effect in online dating: Effects of text-based CMC vs. videoconferencing before meeting face-to-face. *Media Psychology* 23, 6 (2020), 820–839.

[5] Baidu Baike. 2021. Soul. Retrieved September 2021 from https://baike.baidu.com/item/soul/19832271?fr=aladdin#reference-[21]-20525405-wrap

[6] Susan B Barnes. 2006. A privacy paradox: Social networking in the United States. *First Monday* (2006).

[7] Sampson Lee Blair and Timothy J Madigan. 2016. Dating attitudes and expectations among young Chinese adults: An examination of gender differences. *The Journal of Chinese Sociology* 3, 1 (2016), 1–19.

[8] Sampson Lee Blair and Timothy J Madigan. 2019. Dating, marriage, and parental approval: An examination of young adults in China. *Social Science Quarterly* 100, 6 (2019), 2351–2368.

[9] Virginia Braun and Victoria Clarke. 2012. *Thematic analysis.* American Psychological Association.

[10] Ángel Castro and Juan Ramón Barrada. 2020. Dating apps and their sociodemographic and psychosocial correlates: A systematic review. *International Journal of Environmental Research and Public Health* 17, 18 (2020), 6500.

[11] Juliet Corbin and Anselm Strauss. 2014. *Basics of qualitative research: Techniques and procedures for developing grounded theory.* Sage publications.

[12] Judith S Donath. 2007. Social Signals in Supernets. *J. of Computer-Mediated Communication* 13, 1 (2007).

[13] Kirk W Duthler. 2006. The politeness of requests made via email and voicemail: Support for the hyperpersonal model. *Journal of computer-mediated communication* 11, 2 (2006), 500–521.

[14] Nicole Ellison, Rebecca Heino, and Jennifer Gibbs. 2017. Managing Impressions Online: Self-Presentation Processes in the Online Dating Environment. *Journal of Computer-Mediated Communication* 11, 2 (07 2017), 415–441. https://doi.org/10.1111/j.1083-6101.2006.00020.x arXiv:https://academic.oup.com/jcmc/article-pdf/11/2/415/22317530/jjcmcom0415.pdf

[15] Dean Esmay. 2004. Usenet Personals: Advice for Straights FAQ (1/3). Retrieved September, 2021 from 11.http://www.faqs.org/faqs/personals/straightfaq/part1

[16] Eli J Finkel, Paul W Eastwick, Benjamin R Karney, Harry T Reis, and Susan Sprecher. 2012. Online dating: A critical analysis from the perspective of psychological science. *Psychological Science in the Public interest* 13, 1 (2012), 3–66.

[17] Andrew T Fiore and Judith S Donath. 2004. Online personals: An overview. In *CHI'04 extended abstracts on Human factors in computing systems.* 1395–1398.

[18] Andrew T Fiore and Judith S Donath. 2005. Homophily in online dating: when do you like someone like yourself?. In *CHI'05 extended abstracts on Human factors in computing systems.* 1371–1374.

[19] Jeana H Frost, Zoe Chance, Michael I Norton, and Dan Ariely. 2008. People are experience goods: Improving online dating with virtual dates. *Journal of Interactive Marketing* 22, 1 (2008), 51–61.

[20] Ateret Gewirtz-Meydan and Liat Ayalon. 2018. Forever young: Visual representations of gender and age in online dating sites for older adults. *Journal of women & aging* 30, 6 (2018), 484–502.

[21] Jennifer L Gibbs, Nicole B Ellison, and Chih-Hui Lai. 2011. First comes love, then comes Google: An investigation of uncertainty reduction strategies and self-disclosure in online dating. *Communication Research* 38, 1 (2011), 70–100.

[22] Ralph Gross and Alessandro Acquisti. 2005. Information revelation and privacy in online social networks. In *Proceedings of the 2005 ACM workshop on Privacy in the electronic society.* 71–80.

[23] Jeffrey T Hancock and Philip J Dunham. 2001. Impression formation in computer-mediated communication revisited: An analysis of the breadth and intensity of impressions. *Communication research* 28, 3 (2001), 325–347.

[24] Jeffrey T Hancock and Catalina L Toma. 2009. Putting your best face forward: The accuracy of online dating photographs. *Journal of Communication* 59, 2 (2009), 367–386.

[25] Changyang He, Lu He, Zhicong Lu, and Bo Li. 2023. Seeking Love and Companionship through Streaming: Unpacking Livestreamer-moderated Senior Matchmaking in China. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI '23).* Association for Computing Machinery, New York, NY, USA, Article 274, 18 pages. https://doi.org/10.1145/3544548.3581195

[26] E-Ling Hsiao et al. 2012. Synchronous and asynchronous communication in an online environment: Faculty experiences and perceptions. *Quarterly review of distance education* 13, 1 (2012), 15.

[27] GP Ingram, M Isabela Enciso, Nathalia Eraso, Maria Jose García, and A Olivera-La Rosa. 2019. Looking for the right swipe: Gender differences in self-presentation on Tinder profiles. *Annu. Rev. Cyberther. Telemed* 149 (2019).

[28] Linda A Jackson and Jin-Liang Wang. 2013. Cultural differences in social networking site use: A comparative study of China and the United States. *Computers in human behavior* 29, 3 (2013), 910–921.

[29] L Crystal Jiang, Natalie N Bazarova, and Jeffrey T Hancock. 2011. The disclosure–intimacy link in computer-mediated communication: An attributional extension of the hyperpersonal model. *Human communication research* 37, 1 (2011), 58–77.

[30] L Crystal Jiang, Natalya N Bazarova, and Jeffrey T Hancock. 2013. From perception to behavior: Disclosure reciprocity and the intensification of intimacy in computer-mediated communication. *Communication Research* 40, 1 (2013), 125–143.

[31] Kyuha Jung, Yoobin Park, Hanwool Kim, and Joonhwan Lee. 2022. Let's Talk @Clubhouse: Exploring Voice-Centered Social Media Platform and Its Opportunities, Challenges, and Design Guidelines. In *Extended Abstracts of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI EA '22).* Association for Computing Machinery, New York, NY, USA, Article 450, 6 pages. https://doi.org/10.1145/3491101.3519774

[32] Ernest Kramer and Charles D Aronovitch. 1970. Voice expression and rated extraversion. *Journal of projective techniques and personality assessment* 34, 5 (1970), 426–427.

[33] Jeffrey H Kuznekoff and Lindsey M Rose. 2013. Communication in multiplayer gaming: Examining player responses to gender cues. *New Media & Society* 15, 4 (2013), 541–556.

[34] Adeline Y Lee and Amy S Bruckman. 2007. Judging you by the company you keep: dating on social networking sites. In *Proceedings of the 2007 international ACM conference on Supporting group work.* 371–378.

[35] Soohyung Lee and Muriel Niederle. 2015. Propose with a rose? Signaling in internet dating markets. *Experimental Economics* 18, 4 (December 2015), 731–755. https://doi.org/10.1007/s10683-014-9425-9

[36] Zhicong Lu, Michelle Annett, Mingming Fan, and Daniel Wigdor. 2019. "I feel it is my responsibility to stream": Streaming and Engaging with Intangible Cultural Heritage through Livestreaming. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland UK) *(CHI '19).* Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3290605.3300459

[37] Zhicong Lu, Michelle Annett, and Daniel Wigdor. 2019. Vicariously Experiencing it all Without Going Outside: A Study of Outdoor Livestreaming in China. *Proc. ACM Hum.-Comput. Interact.* 3, CSCW, Article 25 (nov 2019), 28 pages. https://doi.org/10.1145/3359127

[38] Zhicong Lu, Haijun Xia, Seongkook Heo, and Daniel Wigdor. 2018. You Watch, You Give, and You Engage: A Study of Live Streaming Practices in China. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3173574.3174040

[39] Christina Masden and W. Keith Edwards. 2015. Understanding the Role of Community in Online Dating. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (Seoul, Republic of Korea) *(CHI '15)*. Association for Computing Machinery, New York, NY, USA, 535–544. https://doi.org/10.1145/2702123.2702417

[40] Julia Naftulin. 2019. Dating apps are turning to message-free meetups and text-only profiles to battle online dating fatigue. Retrieved September, 2021 from https://www.insider.com/dating-apps-the-league-bounce-old-school-methods-2019-12

[41] Flavia Oprea, Horia Virgil Velciu, Andrei Serițan, and Laurențiu Olteanu. 2020. Voice Based Dating Application. In *2020 19th RoEduNet Conference: Networking in Education and Research (RoEduNet)*. IEEE, 1–6.

[42] Aditi Paul, Saifuddin Ahmed, and Karolina Zaluski. 2022. Does Online Dating Promotion Vary Across Cultures? A Cross-cultural Analysis of Homepage Advertisements of Online Dating Services in 51 Countries. *Journal of Creative Communications* 17, 2 (2022), 179–198.

[43] Altman Yuzhu Peng. 2021. Gender and the privacy paradox in Chinese college students' locative dating communication. *Global Media and China* 6, 2 (2021), 225–240.

[44] Artemio Ramirez, Erin M Sumner, Christina Fleuriet, and Megan Cole. 2015. When online dating partners meet offline: The effect of modality switching on relational communication between online daters. *Journal of Computer-Mediated Communication* 20, 1 (2015), 99–114.

[45] Larry D Rosen, Nancy A Cheever, Cheyenne Cummings, and Julie Felt. 2008. The impact of emotionality and self-disclosure on online dating versus traditional dating. *Computers in Human Behavior* 24, 5 (2008), 2124–2157.

[46] Barış Sevi, Tuğçe Aral, and Terry Eskenazi. 2018. Exploring the hook-up app: Low sexual disgust and high sociosexuality predict motivation to use Tinder for casual sex. *Personality and Individual Differences* 133 (2018), 17–20.

[47] Chenxinran Shen, Zhicong Lu, Travis Faas, and Daniel Wigdor. 2021. The Labor of Fun: Understanding the Social Relationships between Gamers and Paid Gaming Teammates in China. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, Article 140, 15 pages. https://doi.org/10.1145/3411764.3445550

[48] Elizabeth Sillence and Chris Baber. 2004. Integrated digital communities: combining web-based interaction with text messaging to develop a system for encouraging group communication and competition. *Interacting with computers* 16, 1 (2004), 93–113.

[49] Aaron Whitman Smith and Maeve Duggan. 2013. *Online dating & relationship*. Pew Research Center Washington, DC.

[50] Gavin Smith. 2013. Dating Services in the US Industry Market Research Report from IBISWorld Has Been Updated. Retrieved September, 2021 from http://www.prweb.com/releases/2013/3/prweb10537905.htm

[51] Randy Jay C Solis and Ka Yee J Wong. 2019. To meet or not to meet? Measuring motivations and risks as predictors of outcomes in the use of mobile dating applications in China. *Chinese Journal of Communication* 12, 2 (2019), 204–223.

[52] Jessica L Spence, Matthew J Hornsey, Eloise M Stephenson, and Kana Imuta. 2022. Is your accent right for the job? A meta-analysis on accent bias in hiring decisions. *Personality and Social Psychology Bulletin* (2022), 01461672221130595.

[53] Michael Spence. 2002. Signaling in retrospect and the informational structure of markets. *American Economic Review* 92, 3 (2002), 434–459.

[54] Brian H Spitzberg and Gregory Hoobler. 2002. Cyberstalking and the technologies of interpersonal terrorism. *New media & society* 4, 1 (2002), 71–92.

[55] Susan Sprecher. 2014. Initial interactions online-text, online-audio, online-video, or face-to-face: Effects of modality on liking, closeness, and other interpersonal outcomes. *Computers in Human Behavior* 31 (2014), 190–197.

[56] Susan Sprecher and Pamela C Regan. 2002. Liking some things (in some people) more than others: Partner preferences in romantic relationships and friendships. *Journal of Social and Personal Relationships* 19, 4 (2002), 463–481.

[57] Ningjing Tang, Lei Tao, Bo Wen, and Zhicong Lu. 2022. Dare to Dream, Dare to Livestream: How E-Commerce Livestreaming Empowers Chinese Rural Women. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 297, 13 pages. https://doi.org/10.1145/3491102.3517634

[58] Catalina Toma. 2006. An Examination of Deceptive Self-Presentation in Online Dating Profiles. (2006).

[59] Catalina L Toma. 2010. Perceptions of trustworthiness online: the role of visual and textual information. In *Proceedings of the 2010 ACM conference on Computer supported cooperative work*. 13–22.

[60] Suvi Uski and Airi Lampinen. 2016. Social norms and self-presentation on social network sites: Profile work in action. *New media & society* 18, 3 (2016), 447–464.

[61] Joseph B Walther. 1996. Computer-mediated communication: Impersonal, interpersonal, and hyperpersonal interaction. *Communication research* 23, 1 (1996), 3–43.

[62] Joseph B Walther, Celeste L Slovacek, and Lisa C Tidwell. 2001. Is a picture worth a thousand words? Photographic images in long-term and short-term computer-mediated communication. *Communication research* 28, 1 (2001), 105–134.

[63] Ye Wang, Zhicong Lu, and Roger Wattenhofer. 2022. Gay Dating on Non-dating Platforms: The Case of Online Dating Activities of Gay Men on a Q&A Platform. *Proc. ACM Hum.-Comput. Interact.* 6, CSCW2, Article 298 (nov 2022), 23 pages. https://doi.org/10.1145/3555189

[64] Janelle Ward. 2017. What are you doing on Tinder? Impression management on a matchmaking mobile app. *Information, Communication & Society* 20, 11 (2017), 1644–1659.

[65] Barry Wellman. 1999. Net-surfers don't ride alone: Virtual communities as communities. *Networks in the global village* (1999).

[66] Martin King Whyte. 1997. The fate of filial obligations in urban China. *The China Journal* 38 (1997), 1–31.

[67] Natsuko Yamada, Yuji Hakoda, Tomoyasu Nakamura, Sanae Shirasawa, and Akiko Kusuhara. 2001. Impression of voice in relation to stereotyped occupational categories. *The Japanese Journal of Psychonomic Science* 20, 1 (2001), 1–12.

[68] Nick Yee, Nicolas Ducheneaut, Les Nelson, and Peter Likarish. 2011. Introverted elves & conscientious gnomes: the expression of personality in world of warcraft. In *Proceedings of the SIGCHI conference on human factors in computing systems*. 753–762.

[69] Yimuliaoran2018. 2018. Soul Experience Analysis Report (in Chinese). Retrieved September 2021 from https://www.jianshu.com/p/b312209d899b

[70] Lotus Zhang, Lucy Jiang, Nicole Washington, Augustina Ao Liu, Jingyao Shao, Adam Fourney, Meredith Ringel Morris, and Leah Findlater. 2021. Social Media through Voice: Synthesized Voice Qualities and Self-presentation. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW1 (2021), 1–21.

[71] Miron Zuckerman, Holley Hodgins, and Kunitate Miyake. 1990. The vocal attractiveness stereotype: Replication and elaboration. *Journal of Nonverbal Behavior* 14, 2 (1990), 97–112.

[72] Douglas Zytko, Nicholas Furlo, Bailey Carlin, and Matthew Archer. 2021. Computer-Mediated Consent to Sex: The Context of Tinder. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW1 (2021), 1–26.