

# Can Fans Build Parasocial Relationships through Idols' Simulated Voice Messages?: A Study of AI Private Call Users' Perceptions, Cognitions, and Behaviors

EUN JEONG KANG, Cornell University, USA

HAESOO KIM\*, Cornell University, USA

HYUNWOO KIM\*, KAIST, Republic of Korea

SUSAN R. FUSSELL, Cornell University, USA

JUHO KIM, KAIST, Republic of Korea

Celebrities have used their imitated voices in media content for various purposes to entertain audiences. However, it is unclear how their fans perceive media content using synthetic celebrity voices. Fans are dedicated to celebrities and actively communicate with them on social media, sometimes resulting in the establishment of parasocial relationships. Focusing on simulated voice messages that mimic celebrities' communication behaviors, we explored how fans form parasocial relationships through synthetic media content. We conducted semi-structured interviews with 15 fans who used AI Private Call, an AI-based voice message service that employed K-Pop idols' synthetic voices. Findings show that the nature of synthetic voices could not provide impressive experiences related to their idols. However, participants believe that simulated voice messages can help make up for the limited opportunities to interact directly with idols. Based on the theoretical framework of parasocial relationship development [92], participants engaged with simulated voice messages, exhibiting various cognitive and behavioral actions that dynamically advanced the relationship. These behaviors include aligning voice content with idols' vocal characteristics, adhering to social norms in interactions, and confirming idols' involvement in content creation to demonstrate a sense of authenticity. Additionally, their communities associated with fandoms affect their experiences throughout the process. Based on the findings, we discuss how simulated voices might impact the relationship between fans and celebrities, highlighting the responsibility of communities and platforms. We also propose design guidelines for creating simulated celebrity voice content that fans can entertain.

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

Additional Key Words and Phrases: parasocial relationship, fan, fandom, voice message, synthetic media

## ACM Reference Format:

Eun Jeong Kang, Haesoo Kim, Hyunwoo Kim, Susan R. Fussell, and Juho Kim. 2025. Can Fans Build Parasocial Relationships through Idols' Simulated Voice Messages?: A Study of AI Private Call Users' Perceptions, Cognitions, and Behaviors. *Proc. ACM Hum.-Comput. Interact.* 9, 2, Article CSCW044 (April 2025), 31 pages. <https://doi.org/10.1145/3711111>

---

\*Both authors contributed equally to this research.

---

Authors' Contact Information: Eun Jeong Kang, [ek646@cornell.edu](mailto:ek646@cornell.edu), Cornell University, Ithaca, NY, USA; Haesoo Kim, [hk778@cornell.edu](mailto:hk778@cornell.edu), Cornell University, Ithaca, NY, USA; Hyunwoo Kim, [khw0726@kaist.ac.kr](mailto:khw0726@kaist.ac.kr), KAIST, Daejeon, Republic of Korea; Susan R. Fussell, [sfussell@cornell.edu](mailto:sfussell@cornell.edu), Cornell University, Ithaca, USA; Juho Kim, [juhokim@kaist.ac.kr](mailto:juhokim@kaist.ac.kr), KAIST, Daejeon, Republic of Korea.



This work is licensed under a Creative Commons Attribution-NoDerivatives 4.0 International License.

© 2025 Copyright held by the owner/author(s).

ACM 2573-0142/2025/4-ARTCSCW044

<https://doi.org/10.1145/3711111>

## 1 Introduction

The applications of AI increasingly open up possibilities for artists to entertain their audiences. For example, Canadian musician Grimes utilizes her synthetic voice to create songs, allowing her fans to create songs with AI voices [45]. Icelandic entertainer Björk used an AI model to generate a song according to the weather [23]. With recent advances in using audio samples to automatically create voice messages that resemble actual persons, celebrities have started explore the potentials of simulated voice messages, or AI-based voice messages simulating celebrities, to increase audience engagement. For instance, live streamers employ synthetic voices to read out donation messages sent by fans who provide financial support while streamers focus on playing games [77].

While such commercially viable products could increase audience engagement, it remains uncertain how AI-based content affects relationships fans have built with celebrities. Fans, who are dedicated to being connected with celebrities, often develop strong parasocial relationships (PSRs) – feelings of intimate, interpersonal-like relationships that arise from one-sided interactions [61]. Based on their attachment to celebrities, fans engage with celebrity-based content or actively interact with celebrities across various social media platforms [92]. However, the use of AI to simulate celebrities in media is new and experimental, creating distinct media experiences. In addition, AI-generated content is often perceived as untrustworthy and inauthentic [73]. Thus, fans might show different user engagement when interacting with AI-generated content of their celebrities. We are interested in how characteristics of celebrity AI-based simulated voice messages provoke fans to develop PSRs.

As confirmed by previous research, the voices of celebrities engage fans [96] and audio media, such as podcasts, can facilitate PSRs between radio hosts and listeners [80]. However, such media research presupposes that audiences heuristically perceive celebrity voices as authentic and genuine. Only few studies indicate that how users use smart speakers using voices of intimate acquaintances could potentially affect their relationships with those acquaintances [10]. However, there is no clear indication of what motivated them for specific behavioral responses that affect relationships. To date, there has been little consideration of the use of voice messages with cloned voices in parasocial relationships, in which users have a high level of engagement with the celebrity whose voice is being cloned.

PSRs evolve over time as audiences engage with celebrities through media exposure [42, 92]. Tukachinsky and Stever's [92] suggest a theoretical model of the developmental processes involved in establishing PSRs. The model identifies individuals' behaviors and cognitions at four stages of the development of PSRs – *initiation, experimentation, intensification, and integration*. As fans move their stage from initiation to integration, they develop a strong bond with their like-minded individuals and celebrities, and actively seek to interact with celebrities. With the lens of this theoretical model, we investigate 1) how fans exhibit thoughts and actions to form PSRs and 2) how fans advance PSRs based on their thoughts and actions.

We are particularly interested in simulated voice messages that emulate the communication behavior of idols. Following Yue and Cheung [12], we define idols as celebrities who are perceived as especially idealistic and lovable figures, and who elicit a considerable degree of passion from their fans. Idols often engage with their fans through CMC channels, thereby creating the appearance of an interpersonal connection. Moreover, idols may strategically employ their communication skills that promote the development of PSRs [76]. For example, idols may communicate intimately with their fans on social media, replying to fans' posts as if they are personal acquaintances.

We conducted semi-structured interviews with 15 fans who had used a service providing simulated voice messages using their favorite idols' voices. The service, called 'AI Private Call' [38], is an AI-based voice message service using K-pop (Korean popular culture) idols' synthetic voices. Fans

could receive voice messages with idols' synthetic voices calling their names according to chosen scenarios. The service was launched specifically for fans through official contracts with idols, and it emulated interpersonal communication with realistic voices. By understanding AI Private Call users' psychological responses to simulated voice messages within the framework of Tukachinsky and Stever's theoretical model, our goal was to understand how these simulated voice messages affected fan-idol relationships, particularly the development of PSRs. Our research questions are:

- RQ1. How do fans perceive voice messages that emulate idols' communication behaviors?
- RQ2. What cognitions and behaviors to establish parasocial relationships do fans exhibit in the experience of idols' simulated voice messages?

From our interviews, we observed that fans hold two perspectives on simulated voice messages based on components comprising the content: 1) the voice of a virtual agent, yet 2) the message portraying how their idol might behave in interpersonal relationships. We also found that fans proceeded with a series of steps in engaging with simulated content of their idols. Fans initially compared the voice message for consistency with their prior knowledge of the idols' voices, interests, and activities. Then, they assessed whether the voice message is appropriate in interpersonal communication based on their social norms. Subsequently, they actively sought information from other channels to understand other fans' experiences with simulated voice messages. To deepen the intimacy through the user experience, fans assessed how their idols were involved in content creation and sought ways to increase their idols' social presence over the content, such as listening to idols' audio content and adjusting interface settings. They also shared experiences of simulated voice messages with their fandoms. Each step in these sequential process was associated with a series of stages to establish PSRs, suggesting that simulated voice messages may contribute to engaging with idols. Based on the findings, we discuss what to consider in designing user experience of synthetic media that simulates celebrities' voices or images, so as to increase fan engagement.

Our contributions include:

- Identification of fans' perceptions of simulated voice messages and how simulated voice messages contribute to building PSRs over multiple stages.
- Implications for designing synthetic media content with celebrity's simulated voices integrated into media culture.

## 2 Related Work

In this section, we review related work in four areas: 1) synthetic media using celebrities' simulated voices, 2) parasocial relationships and fans, 3) developing relationship through listening, and 4) user experiences of simulated voice messages.

### 2.1 Synthetic Media Using Celebrities' Simulated Voices

Celebrities have begun producing synthetic media employing their voices artificially generated by computers. With only a limited number of audio samples, they can generate content using synthesized voices, including songs and audio-books. Due to its simplicity and customizability in production, numerous business opportunities have emerged to commercialize celebrities' voices [16]. AI-based simulated voices can also be used for interactive and personalized experiences without the direct involvement of celebrities. For example, a language education material utilized AI synthetic voices modeled after BTS members, a K-Pop idol group, to assist foreign fans individually in learning Korean language [100, 103]. Similarly, streamers can utilize their simulated voice to acknowledge fans for donation by calling out viewers' names in live streams [77].

Most studies on AI-based content for high-profile individuals have considered ethical and legal implications to protect their publicity rights and privacy from artificially created content [20, 58, 74].

Another research suggests that AI-generated content is usually perceived as less authentic. For example, advertisements with AI-generated content potentially reduce the intention to purchase the advertised products [73]. However, relational outcomes of engaging with content that imitates public figures have not been determined. Relationship is important in perceiving one's partner as it enables individuals to offer affection, support, and communication to their partner [99]. Audiences who maintain relationships (what is termed as 'parasocial relationship') with celebrities often exhibit specific behaviors when engaging with celebrity content [92]. In turn, celebrities often manage their relationship with fans to maintain their media impacts by interacting with fans on social media [39]. Celebrity's authentic voice has been shown to increase user engagement in traditional media [96], but synthetically generated voice may offer different experiences for fans.

We aim to explore how fans, who frequently form parasocial relationships with celebrities, engage with synthetic media, particularly voice messages, and how their behaviors can foster relationships. We believe that this study, grounded in a real-world application, will inform the development of synthetic celebrity content to enhance fan engagement.

## 2.2 Parasocial Relationships and Fans

When audiences recognize and interact with human representations appearing in the media, parasocial relationship (PSR) can be produced in their imagined interactions [26]. PSR refers to one-sided relationship where the audience feels involved with a celebrity like interpersonal relationship, but the celebrity is unlikely know who they are [14, 26, 92].

PSR does not uniformly manifest to the same degree in every user's experiences. PSRs can evolve as continuum of encountering their media figures in media as the level of engagement with celebrities grows [26, 42, 92]. Tukachinsky and Stever [92] have conceptualized the stages of PSR as initialization, experimentation, intensification, and integration. Each stage involves distinct cognitive, affective, and behavioral characteristics (Table 1). The progression through these stages can vary depending on the quality and depth of interactions experienced by the audience. Furthermore, the audience may exhibit behaviors associated with multiple stages simultaneously, such as demonstrating both intensification and integration behaviors concurrently.

**2.2.1 Initiation.** Audiences form initial impressions of celebrities based on characteristics or physical appearance portrayed through the media. Through these media exposures, audiences are initially drawn to and begin to pay attention to celebrities. Due to the lack of information about celebrities, audiences examine traits of these figures, evaluate them, and engage in social comparisons with others [92].

**2.2.2 Experimentation.** Some audience members, or fans may actively strive to cultivate deeper relationships with celebrities by gathering information about them and seeking to gain a better understanding of them. In comparison with the 'audience', 'fan' is broadly defined as a person who actively seeks to engage with media content rather than casually watch, listen, or engage [24, 35]. In this stage, fans explore a higher volume of media content associated with celebrities to reinforce engagement and feelings of connection. They seek consistency and coherence across different media platforms, aiming to align their perceptions of their celebrities with characters of their celebrities exposed in media content [92].

**2.2.3 Intensification.** Some fans can actively seek and maintain relationships with celebrities, striving to further enhance their connections. They perceive celebrities as parental figures or friends, mirroring the characteristics of interpersonal relationships. Additionally, they re-watch content and seek additional sources. Celebrities' self-disclosures and perceived reciprocity play crucial roles in fostering the development of these relationships. In this regard, celebrities' strategic

communication through social media can significantly influence individuals. On the other hand, incidents that violate a fan's expectation or disrupt their perceived relationship dynamics may impede the development of PSRs [92].

**2.2.4 Integration.** Finally, some fans can integrate characteristics of celebrities with their feelings of self. Specifically, fans may feel they have a special connection with fandoms that consist of like-minded individuals. Fandoms and celebrities contribute to the formation of an individual fan's self-concept, mood, and identity. Additionally, through deep engagement, fan can gain more autonomy in interpreting celebrities, enabling a subjective understanding of a celebrity's personality and feelings [92].

The model helps assess the extent of user engagement in the media based on behaviors and cognition associated with stages, and identifies factors that move to a deeper PSR stage. It can also enhance our understanding of the media's influence on self-concept based on media effects or relational outcomes observed at each stage. Using the lens of Tukachinsky and Stever's [92] theoretical model of PSRs, we investigate what fans' behaviors and cognitive responses to the simulated voice messages are related to PSRs and how they affect development of PSRs.

### 2.3 Developing Relationship through Listening

A significant amount of research indicates that humans can evoke parasocial responses by interacting with digital entities like chatbots [66] and virtual influencers [13, 49, 50, 54, 87]. The characteristics these entities possess, such as visual appeal and human-like traits, enable people to form social connections. Comparatively, the effect of voices, especially synthetic voices, on relational interactions is under-explored despite its extensive potential as media content.

Listening behaviors are everyday communicative practices that constitute a relational process [71]. Prior research indicates that listening practices can foster and sustain relationships, and the attitudes and behaviors of either listeners or speakers can positively impact parasocial relationships [1, 8, 60, 71, 80, 95]. Relational listening, a listening behavior that motivates listeners to focus on speakers' emotions, can positively affect parasocial relationships [95]. A podcast host's interactions, such as sharing their personal lives, can also positively influence relational establishment [80]. This also includes other type of content which creators create as sensory activities (e.g., ASMR) to elicit affective experiences [65, 82].

Another line of research suggests that the relationship established with the speaker allows listeners to pay attention to other audio content using the speaker's voice [8, 10, 32]. Familiar voices, such as one's own voice or a friend's voice, are often perceived as more engaging and trustworthy, which can be utilized for various purposes, from intervention to habit formation [10, 32, 40]. Chan et al. [10] examined how the voices of close friends and relatives in virtual assistants can enhance engagement, persuasion, and trustworthiness. On the other hand, Kim et al. [40] showed that a self-voice alarm can effectively serve as a daily reminder to encourage users to form habits.

The research series highlights that familiar, genuine voices and computer-mediated interactions between speakers and listeners can help foster relationships. However, synthetic celebrity voices are under a different circumstance, in that the listener is likely to engage in the content without the speaker's perceived presence.

### 2.4 User Experience of Simulated Voice Messages

We focus on 'simulated voice messages', which use synthetic voices to impersonate target speakers and convey information without verbal interactions. In line with the significant advancement of synthetic voice technologies in acoustic quality, HCI research suggests designing synthetic voice content from user-centered perspectives to reduce negative experiences like the uncanny

Stages	Cognitive Level	Affective Level	Behavioral Level
Initiation	Individuals explore the character, critically assess him or her, and make social comparison between others and figures.	Individuals are attracted to physical characteristics of figures which can readily be recognized at first sight.	Individuals seek ways to increase media exposure in order to lessen uncertain information about figures and explore potential relationships.
Experimentation	Individuals confirm consistency of figures' personality across channels, developing stronger opinions than prior stage.	Individuals develop homophily with character more than initiation stage.	Individuals engage in multiple interactions and comparisons across various media channels to explore additional information about figures.
Intensification	Individuals frequently ponder media figures in their lives, imagining interactions as a part of interpersonal relationships.	Individuals develop strong emotional connections toward the media figures.	Individuals seek additional context in media channels where they can deepen relationships (e.g., social media), engaging in discussions with others about the character.
Integration	Individuals perceive media figures' character or their fandoms as their personal identity.	Individuals form deep intimacy with media figures, experiencing a profound and special bond.	Individuals invest significant time on parasocial relationship activities (e.g., writing fan-fiction) and socialize with like-minded individuals.

Table 1. The theoretical model of the development of parasocial relationships suggested by Tukachinsky and Stever [92]

valley [3, 4]. First, synthetic voice should consider voice expressiveness humans possess in various situations [3, 64]. Noufi et al. [67] suggested that humans formalize their voices considering vocal persona adjusted by external and internal factors such as self-expressions, physical environment, and how they perceive listeners. Second, all listeners can listen to the content in different contexts with different needs and subjective preferences [9, 19, 48]. Cambre et al. [9] demonstrated that listeners have different preferences in listening to long-form content, also indicating that synthetic voices can be preferred if they are designed considering the listening purposes. Do et al. [19] highlighted that trustworthiness of a voice can be perceived differently based on the gender. In relation to simulated voice content, recent research on simulated voice content has focused on designing these voices responsibly to protect against deceptive tactics used in scams and harming target speakers [33, 51].

Based on these explorations, we can anticipate that simulated voice messages may elicit interactions to contribute to relationships. However, it is unclear that how unique relationships and communities associated with celebrities, where fans are deeply embedded, affect media engagement

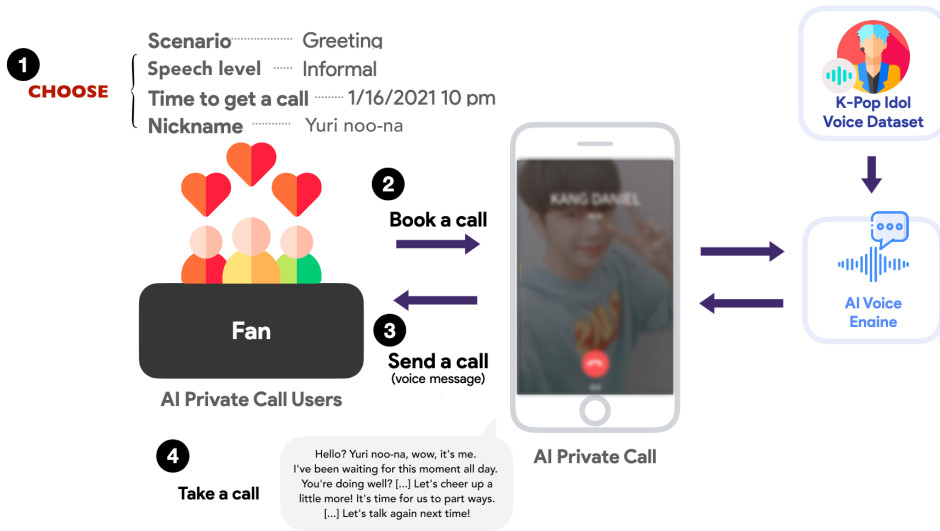


Fig. 1. A user's interaction process of AI Private Call, 1) A user chose a scenario, speech level, time to get a call, and a name in the interface, 2) When all options were chosen, a prompt said, 'The call is booked.' 3) When the designated time arrived, the user's phone rang with the photograph of an idol on the screen. 4) When the user took a call, they could listen to a voice message using an idol's synthetic voice. For example, when the user chose the following options in the application with an idol (Kang Daniel) subscription ticket: 1) scenario: greeting, 2) speech level: informal, and 3) name: Yuri noo-na, the incoming call the user received was as follow: "(ringtone) Hello? Yuri noo-na?, wow, it's me. I've been waiting for this moment all day. How's going? Yuri noo-na, Cheer up. I believe you'll do a great job. I am rooting for you! It's time for us to say goodbye. That's too bad. Let's talk again next time!"

and how parasocial relationships are built through specific behaviors and cognition. Thus, in this study we aim to investigate from fans' behaviors and cognition in listening to idols' simulated voice messages.

### 3 CONTEXT OF RESEARCH: AI PRIVATE CALL

We describe the concept of AI Private Call and how fans use it to listen to simulated voice messages in the service.

#### 3.1 Concept of the Service

'AI Private Call' was an AI-based voice messaging service that simulated a personal phone call from a K-Pop idol. It was offered as an exclusive feature for paid subscribers on the fandom community mobile application 'Universe' which was discontinued in 2023. Subscribers could access both 'AI Private Call' and 'Private Message' – a non-AI based text message service where an idol could actually send text messages to fans [102]. Each idol's paid subscription plan cost 7,900 KRW (approximately 6.65 USD) monthly, with additional fees per month for subscribing to multiple idols.

## 3.2 Simulated Voice Message

The user could receive a voice message in Korean that simulated a personal call from their chosen idol when they chose a date, a speech level, and a scenario through the application. This message simulated the experience of receiving a personal call from an idol without any actual verbal interaction, using the chosen idol's synthetic voice (Figure 1).

**3.2.1 Synthetic voice.** The synthetic voice was modeled after an idol's actual voice, based on an AI-based text-to-speech (TTS) model trained with their voice data [38]. The application company (NCsoft) officially contracted with idols, then utilized their voices to create synthetic voices.

**3.2.2 Script.** The script of the voice messages followed a plot based on a selected scenario, broadly consisting of 1) a greeting, 2) a chat related to the selected scenario, and 3) a wrap-up. This script could be modified based on the chosen options for customization in the interface. Each idol's voice message contained different content while maintaining a consistent plot. This allowed users to enjoy subtly varying scripts across different idols' voice messages.

**3.2.3 Interaction on interface for customizing voice message.** As illustrated in Figure 1, the user could adjust options for message customization in the interface; There were four options that a user should choose before completing booking a call: 1) the date and time of a call, 2) the idol's speech level (formal, informal, any – random chosen), 3) a call's scenario (morning call, cheering up, birthday celebration, and daily chat), and 4) the name that a user wishes to be called by (A detail of the option is provided in Appendix A). If the user wanted to listen to voice messages again or miss a call, they could listen to the voice messages again.

## 4 Methodology

We conducted semi-structured interviews with 15 people with experience using AI Private Call, asking them to reflect on how they used it and how it affected their relationships with idols. The study protocol was approved by the institutional review board (IRB). The study was conducted in August 2021, before the service was terminated in June 2022<sup>1</sup>

### 4.1 Participants

Interviewees were recruited through ads in K-pop fan communities, university community channels, and social media platforms (Twitter, Facebook). For K-pop fan communities, we posted ads on idols' public community boards (DC inside [17]) and public fandom chats. Two participants were recruited by snowball sampling [27] from interviewees' fan networks. All participants were native Korean speakers since AI Private Call was only provided in Korean. As AI Private Call was available to users with a paid plan on the platform, we targeted the fans who had subscribed to the paid plan for more than three months to ensure their familiarity with AI Private Call. We asked applicants to submit a pre-survey to gather information about their background of using the service and to confirm their status as genuine fans of the owners (idols) of the artificial voices. The pre-survey included questions about the types of fan activities they had participated in, the frequency of AI Private Call usage, and their satisfaction with the service. During the recruitment process, we

<sup>1</sup>The service was discontinued before the mobile application 'Universe' was deactivated in 2023.



PID	Gender	Age	The number of idols & their genders participants subscribed to	Reported period that participants maintained fan activities <sup>a</sup>	Frequency of using AI Private Call
1	Chose not to disclose	24	1 (M)	-7 months	Less than one time in a week
2	F	23	1 (M)	-2 years	Everyday
3	M	36	4 (F)*	-4 months	Less than one time in a week
4	M	20	1 (F)*	-5 months	3-4 times in a week
5	M	34	4 (F)*	-5 months	Less than one time in a week
6	F	20	4 (F)*	-5 months	Less than one time in a week
7	F	30	1 (M)	-5 years	Less than 3 times
8	F	22	1 (M)	-2 years	Less than 3 times
9	F	29	2 (M)	Chose not to disclose	Less than 3 times
10	F	27	3 (M)	-6 months	Less than 3 times
11	F	25	1 (M)	-2 years	Less than 3 times
12	M	26	4 (F)*	-4 months	Less than 3 times
13	F	20	1 (F)	-2 years	Less than 3 times
14	M	30	4 (F)*	-4 months	Less than 3 times
15	M	27	4 (F)*	-5 months	Less than 3 times

\* They were subscribing to the same idol group  
<sup>a</sup> Fan activities include listening to music, purchasing idol merchandise, participating in fandom communities, and following idols' social media updates.

Table 2. Summary of interview participants.

aimed to balance the gender of the participants and the gender of the idols they subscribed to, as fans of male idol groups and female idol groups might have different interactions with their idols.

Table 2 shows the demographic information for the participants (8 female, 6 male, 1 chose not to disclose, ages = 20 – 36). Each participant subscribed to an average of 2.2 idols for an average of 4.2 months (Max= 7.0 months<sup>2</sup>). Four participants answered they were using AI Private Call less than once a week. Two people stated that they were using the service regularly. On the other hand, nine participants responded that they used AI Private Call less than three times throughout their subscription period. Participants also indicated that they frequently engaged in fan activities by listening to their idols' songs (*Mean: 4.67, Median: 5.0*), subscribing to a paid plan (*Mean: 4.60, Median: 5.0*), purchasing idol-related items (e.g., merchandise) (*Mean: 4.33, Median: 5.0*), following their idols' social media (*Mean: 4.73, Median: 5.0*), and engaging in fan activities in online communities (*Mean: 4.67, Median: 5.0*). All questions were measured using five-point Likert scales.

## 4.2 Interview Protocol

The interviews were conducted over Zoom during August 2021, when AI Private Call was still available. Each interview lasted approximately 60-90 minutes, and the participants were compensated with 15,000 KRW (approximately 13.0 USD). We first asked our participants about their usual fan activities to understand whether and how they had built parasocial relationships with idols through other media channels. We also asked about their previous experience about AI-powered apps. We then asked participants to share their specific experiences with AI Private Call and their motivations for using the service.

We organized the interview questions around five themes: (1) the participant's background as a fan; (2) how they used the service; (3) positive and negative experiences in using the service; (4) predicted influences of AI Private Call on their fan-celebrity relationships; (5) differences between AI Private Call and other fan-celebrity communication services (e.g., live-streaming, social text message) (Appendix B). We designed background questions (1) based on their responses in the pre-survey, which ask about their fan activities to facilitate conversation with them. The interview questions in categories (2) and (3) aimed to collect participants' impressions and past interactions with the service. If participants had discontinued the use of the service, we inquired about their

<sup>2</sup>The service was launched on February 2021

reasons for doing so. Regarding theme (4), our goal was to probe the emotional reactions triggered by the service, particularly how it evoked feelings of connection with their idols in the context of parasocial relationships [44]. Questions under theme (5) sought to understand how current services fulfill the needs of fans in parasocial relationships compared to AI Private Call.

During the interviews, we prompted participants to reflect on their past experiences and the motivations for their responses to the service. We also asked to share about the context of their fan activities that might affect their reactions, as follow-up questions for each question.

### 4.3 Data Analysis

We conducted a thematic analysis using a combination of closed and open coding to analyze the interviews [7, 79]. The interview records were transcribed using ClovaNote [62], and processed to remove all personally identifiable information. First, three authors inferred codes inductively, then understood links between relevance codes using in-vivo coding, then created high-level themes using axial coding. During this process, we identified underlying themes by collaboratively clustering the codes using affinity diagramming and refining these codes through discussion to reach consensus. We defined *perception* as ‘user’s meaningful interpretation of sensory input’, *cognition* as ‘user’s cognitive process that involves assimilating information and applying it in following judgment’, and *behavior* as ‘user’s interactive activities as mental representations to generate external outcomes’ [25, 86].

Then, the first author revisited the codes assigned to the themes, considering which aspects of the simulated voice messages were presented within the codes as either ‘voice’ or ‘message.’ Codes categorized under ‘voice’ were related to vocal characteristics such as intonation, tone, and voice pitch. Codes related to narrated content were categorized under ‘message.’ In this process, the first author assessed whether the codes associated with each aspect were coherent and consistent. For instance, if there is inconsistency between codes under ‘voice’ and codes under ‘message’ in a theme, the author created a newly created sub-theme to assign them to clearly understand if their behaviors or cognition presented in a code are influenced by specific components.

As a final stage, guided by the four stages of theoretical model of parasocial relationship development [92], we examined how the themes and sub-themes are relevant with specific stages. We associated themes with cognitive and behavioral patterns that fans may exhibit towards media content at specific stages. For example, for the theme ‘Sharing content experiences with the fandom (Section 5.2.6),’ we aligned this theme with the ‘integration’ stage. This categorization was based on the observation that the codes within the theme reflected similar responses to those of fans who are deeply engaged with celebrities, introduced in the model. In iterating the process to finalize the categorization, we also discussed to understand the behaviors and cognitive experiences that fans encounter exclusively through simulated voice messages, which are not evident in other media channels (e.g., social media, live-streaming). The representative quotes were originally in Korean and were translated into English using DeepL [18] and Papago [63] after finalizing the themes.

## 5 Findings

In this section, we first introduce fans’ perceptions of simulated voice messages provided by AI Private Call (Section 5.1), then we present six themes that fans process as cognition and behaviors to establish parasocial relationships (Section 5.2).

## 5.1 Perceptions of Simulated Voice Message Emulating Idol's Communication Behaviors (RQ1)

Participants perceived synthetic voice message as a combination of two distinct components – synthetic voice and message content – and they had different perception on each component. Notably, although they recognized that synthetic voices do not give any impression about their idols, they reasonably found the potential to create ‘illusion of intimacy’ with idols, if the message content and its interactions are engaging and tailored to their needs.

*5.1.1 Voice of a Virtual Agent.* Even if simulated voice messages mimic a celebrity's voice, they are unlikely to be perceived as genuine due to their quality compared to actual voices and the information indicating they are AI-generated things. All participants, except one (P9), knew that the voice of the content as being artificially created, because the service name (AI Private Call) implied that the voice is artificially generated and the voice did not show natural vocal characteristics and sound expressivity typically associated with human speech. Plus, as they were already used to hearing their idol voices through other media, they immediately noticed something off, describing the voices as the ‘uncanny valley’. Although they were aware that idols provided voice audio recordings as data for training AI models through official contracts, they often referred to the subject of the simulated voice messages as ‘**Fake [idol's name]**’, ‘**robot**’, ‘**machine**’, and ‘**the AI**’ indicating non-human objects. They also described their listening behaviors as ‘**labor**’, ‘**meaningless work**’.

*“Why should I listen to this **robot**'s voice? I don't understand why I should go through this **labor** to listen to this fake voice. I would just choose listening to [P7 idol's name]' voice through V-app that has a voice-only radio feature.” (P7, F)*

In particular for P11 and P7, making fabricated voices utilizing idols for fan services seemed to deprive them of opportunities to interact with their idols. P11 thought that simulating their idol's real voice is not acceptable, unlike cartoon avatars with idols [21] and humanoid virtual influencers [13] which are presented as separate entities from the idols. She thought that simulated content should not replace what idols currently do for fans and their careers.

*“I object to virtualizing [P11 idol's name]<sup>3</sup> if it is an exact replica. [...] If the virtualization occurs, I wonder why it should be used for communication, given the existence of living idols. Nowadays, there are idols who are virtual characters, but those cases are different. **Why do they (the companies) want to replicate and create it (synthetic voice) with idols who have been doing well in their areas? I haven't become a fan of those virtuals.**” (P11, F)*

*5.1.2 Message that My Idol could Say in Interpersonal Communication.* With the message content simulating interpersonal communication behaviors, participants imagined scenarios where their idols might act in various interpersonal situations. Participants (P2, P4, P5, P12) perceived that the message content of simulated voice messages could fulfill their desire to personally interact with their idols, given the limited opportunities for fans to engage in direct communication with idols.

*“When I first heard about this (AI Private Call), I was curious, like ‘If I subscribe to all of them (idols' group), **Would** it feel like we're talking on the phone?’, then I was so excited. So after its (AI Private Call) launch, [...] I immediately subscribed to the paid plan and listened to them, but the voice sounded stiff.” (P12, M)*

<sup>3</sup>We used ‘[ ]’ mark in de-identifying idols' names mentioned by participants. To describe demonstrative pronouns or adjectives in quotes, we used ‘( )’ mark.

They believed that personal conversation with their idols might be impossible in their lives. This is why P2, P4, and P5 listened to the simulated voice messages despite the service's poor performance. Even it is rare to get a chance to talk with idols for a few minutes unless they win a ticket for video calls through the purchase of physical music albums. P4 stated that *"I know they (idols) usually don't have much time because they're popular, but (AI Private Call) feels like they're making time to talk for me."* Thus, they thought that the message content which emulated communication behaviors could let them imagine the situation.

However, simulated voice messages were limited in providing immersive experience, as they only enabled listening. Interpersonal communication requires a two-way conversation where people who know each other exchange utterances sequentially (P9). Consequently, participants felt dispirited (P1, P3, P10).

*"I assume the service was initially designed to offer the primary feature of allowing users to listen to their favorite idols' voices. However, for fans, it (Behavior for listening to idols' voices) is not just about hearing the voice; it's about understanding the emotions the idol wants to share in their daily life. Thus, fans can feel what the person is upset about, or why they're in a bad mood. I think if the content can make them feel a little bit more like that, the content will be a little bit more successful." (P10, F)*

Several participants (P3, P12, P13) mentioned the movie 'Her' [37] where an AI agent forms a social bond with human, suggesting that engaging in reciprocal verbal communication with an agent simulating idols could potentially foster relationships in the absence of direct interaction with their idols. Similarly, P8 stated that simulating communication behaviors like AI Private Call may give benefits to fans whose idols do not actively engage with their fans. If idols actively update their status on social media, anyone may be able to get informed from them directly. As K-Pop idols commonly interact with their fans over social media, participants who were already exposed to various idol content through YouTube videos and social media, did not need less dynamic content that only involved listening.

*"I think this (satisfaction on AI Private Call) may vary between different idol members. I guess that some idols could be replaced by the AIs, and if they (idols) don't often communicate with fans, there would be people who would be satisfied with their AIs. I think the frequency of idol communication might have a slight impact on the activation of the use of AI. But I have enough content about my idols even if I don't have to use this AI." (P8, F)*

From the perspective of certain participants (P4, P5, P6, P9, P10, P13), the utilization of idols' simulated voice to create idols' content could potentially lead to divergence in media content, in addition to AI Private Call. For fans, simulated voice messages are another form of celebrity content, similar to radio segments featuring the actual voices of celebrities. Thus, these messages could provide additional opportunities for them to engage with their idols.

*"I just feel like I have one more piece of content to enjoy, it is the closest experience I've ever gotten to just watching a YouTube video. [...] For me, I think it's just like any other radio content or anything like that, so obviously, if I listen to it more, I will develop a stronger emotional connection to idols." (P5, M)*

## 5.2 Cognitions and Behaviors, and Parasocial Relationship Development (RQ2)

In this section, we introduce participants' cognition and behaviors exhibited in relation to PSR establishment when they interact with simulated voice messages of AI Private Call. Findings show that fans first heuristically examine content, then recognize content as the production of their idols in establishing PSRs (Section 2). We also found unique reactions to simulated voice messages that

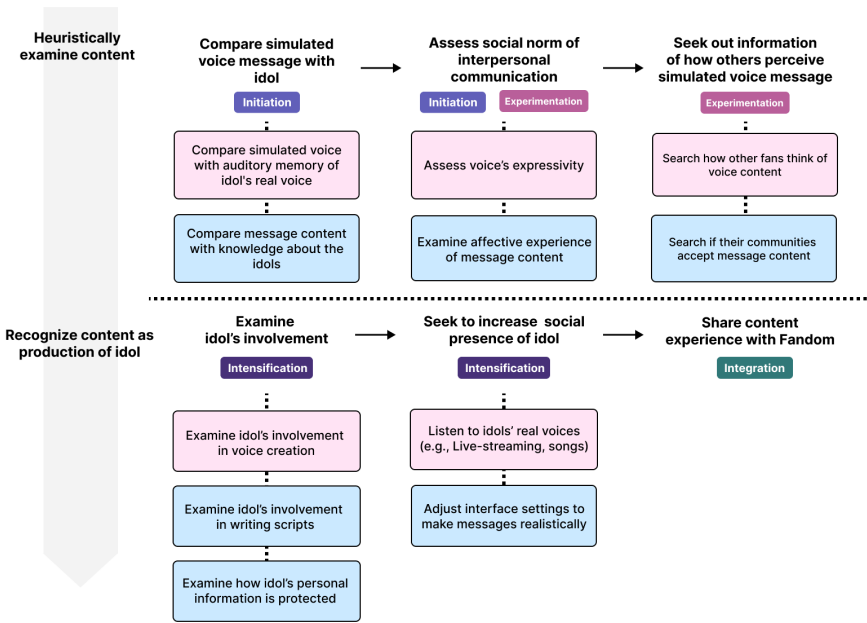


Fig. 2. Parasocial relationship (PSR) developmental process that fans process with simulated voice message. From the top row, each boldface illustrates the main process of responding to simulated voice messages by integrating perception, cognition, and behaviors. Each stage is related to how fans engage with simulated voice messages to foster PSRs. At each stage, fans hold distinctive perception of the simulated voice messages with two content parts: voice (light-pink) and message (light-blue).

differed from the usual responses to authentic celebrity content. Within each theme, we delineated how participants distinctively react to each the voice and message content.

5.2.1 [Initiation] *Comparing Simulated Voice Messages with Knowledge about My Idol.* When audiences first encounter media characters, they critically assess them and use heuristics to gauge their potential for emotional attachment. (initiation stage) [92]. With simulated voice messages, participants' reactions were close to technical evaluation, indicating that they consistently evaluated the extent to which the simulated voice message captured their idols' characteristics, such as speech patterns, based on their knowledge of their idols. In particular, as fans are well informed about their idols, they wanted to interpret the experience based on their knowledge, rather than treating them as a new instance. They sought to achieve consistency between idols' characteristics and simulated voice messages. When they experienced misalignment, they stated that the voice messages lacked distinctiveness associated with their idols.

**Voice content.** Nine participants (P1, P2, P3, P4, P7, P8, P11, P13, P15) characterized idols' vocal personas. They recalled how idols managed their vocal characteristics, such as pitch, perceived emotions, and para-linguistic cues. This was mostly based on sources from online channels, such as YouTube videos and live-streaming content, where they could listen to idols' real voices. Even they speculated how their idols might ideally talk based on typing quirks (e.g., suffixing text with ya in Korean:ㅇㄹ) idols usually used on social media posts. They expected these characteristics to be present in cloned voices.

*“As [P7’s idol name] voice is lower-pitched than others, he usually raises his tone to sound brighter for the fans. [...] But it (synthetic voice) doesn’t contain the characteristic at all. I think it should have something like an intonation. For example, [P7’s idol name] voice and this robot’s voice are different when it says, ‘Did you have a meal?’ This robot sounded to follow a rule to maintain a high tone.” (P7, F)*

**Message content.** Participants (P1, P3, P7, P10, P12, P15) felt that message content should be consistent with how their idols usually communicate with their fans. They synchronized the content of messages with their knowledge such as idols’ preference and topics usually shared by them on social media. Fans typically monitor their favorite celebrities’ activities updated on social media, sharing their information updates with other fans [15, 39]. Additionally, each idol offers unique content that can attract fans. Such information is used as a reference when they listened to simulated voice messages.

*“The [message] content was something they (P15’s idol group) could say in their real life, but since we’re fans, we know their schedules. For example, they (P15’s idol group) had a schedule for a radio show, but it (AI Private Call) just said like ‘I was late for something.’ I felt a gap from such a mismatch.” (P15, M)*

As P15 stated, participants stated that their idols have varied schedules for actively interacting with fans, such as midnight after finishing work, and this information is given by fandom communities or idols by themselves via social media. But the service was open to use regardless of being consistent with the target speakers.

5.2.2 [Initiation, Experimentation] *Assess Social norm of Interpersonal Communication.* Physical or attitudinal attractiveness of characters lead fans to build parasocial relationships [92]. Media experiences allow fans to perceive a sense of playfulness in characters, which encourages them to engage more deeply with them [92]. Similarly, participants were interested in assessing the playfulness and immersion of the content, rather than just the imitation of the idols. In this instance, they examined how simulated voice messages adhere to social norms of interpersonal communication, thus creating engaging experiences that appeal continued use.

**Voice content.** Participants expected that the voice in simulated voice messages would have the same expressiveness they encountered in their various interpersonal communication (P8, P10, P13, P15). They mentioned that the synthetic voice should have adjusted its vocal characteristics such as tone of the voices or para-linguistic cues to adapt to various situations. Their statements echo Aylett et al.’s perspective of voice design [3]. When there was a mismatch between expected voice in a specific situation and synthetic voice of the content, the voice content could make it sound ‘similar but soulless’ and ‘unresponsive’.

*“Since it’s artificial, it’s not like we’re genuinely communicating with each other. But the quality didn’t reach the level that could provide a sense of realism or immersion. It would have appeared more authentic if it had been a recorded voice... They (AI Private Call) claimed the voice was artificially generated, but it felt like I was talking to a wall, as if they were reading a Korean book. Even it sounded too soulless to consider it as a real voice.” (P15, M)*

**Message content.** In the message content, general social norms the speaker usually follows in close relationship are expected. In dyadic conversation, the speaker adjusts their expressions considering if the listener understands the conversation, reactions, and interests. And sometimes they change topics (P1, P2, P3, P7, P8, P9, P10, P11, P12, P13, P14). These behaviors exhibited in interpersonal relationships are mutually inclusive, propelling the relationships forward. Especially

through telephone, people feel gratification from a sense of inclusion, influential impact, and pleasure derived from social interactions [68]. In this manner, participants used their social norms to evaluate the message content, anticipating it to be warm, welcoming, and polite. They also anticipated more personalized topics tailored to their interests that people in close relationships could share.

*"I believe it would have been a lot different if they (AI voice) had waited for me to speak instead of just talking and cutting me off in this one-sided way." (P1, Does not wish to answer)*

Personal conversation with idols would be exclusive and memorable even if they just say, such as 'Happy birthday', but the same utterance spoken by synthetic voice was undervalued by participants (P5, P9, P13). As they became more accustomed to listening to the content, they increasingly sought impressive and unique experiences over what idols could deliver with their own voices from simulated voice messages, whereas they were willing to watch videos or recordings of their idols repetitively.

*"When I used this service for the first time, I was excited and really into it when I heard her voice over the phone. It felt as if she was speaking directly over the phone. However, as I continued using the service, the initial excitement faded, and it began to feel more like listening to a recording on YouTube. Due to the repetition of the same experience, it's inevitable that the sense of novelty diminished. Initially, I attempted to immerse myself in the experience. After listening to it about three times, I found it increasingly difficult to do so. Anyway, it is a machine after all..." (P5, M)*

According to P5, interpersonal conversations are dynamic and cover a range of topics. He expected simulated voice messages to reflect this complexity and unpredictability, rather than following user-selected patterns. Although the service's simulated voice messages offered a slightly different script each time, he found them unentertaining due to their same pattern: a greeting, a chat, then a wrap-up.

**5.2.3 [Experimentation] Seeking out Information in Other Channels.** Within participants' communities, or fandoms, they searched information about how others perceived simulated voice messages or if others also got the same impression as them. Nine participants mentioned that they read reviews of the service on community boards, heard feedback from other fans, or shared their own thoughts when asked by peers (P1, P3, P7, P8, P10, P11, P12, P13, P14). In addition, P9 stated that she received questions from other fans who were interested but hesitant to use the service. Their behaviors mirrored those of audiences in the 'experimentation' stage of PSRs, who strive for achieving consistency in celebrities' identities across contexts and reducing uncertainty about information [92].

**Voice content.** Six participants sought out the opinions (P1, P3, P7, P10, P12, P13) of other fans in media channels they usually interact with or friends who have relevant knowledge. The experiences of others have influenced them to form informed beliefs, opinions, and emotions. They expected other fans would share same impressions of the service as them.

*"When I used AI Private Call for the first time, I was like, 'Is this really what it's like?' then I searched for it and I read all relevant posts, then I was like, 'Oh, I was not wrong. People were feeling in the same way.'" (P1, Chose not to disclose)*

Using synthetic voices of idols sometimes caused confusion in their interactions with idols on other social media channels (P7, P11, P13). This is because they suspected that AI was already being invisibly used in other media. They mentioned that some users wondered whether their idols on social media were also being replaced by AI agents. P7 mentioned that they sought confirmation

by asking confirmation to their idols in Private Message, an affiliated service where idols send text messages. In this case, fans asked their idol directly through the service if he was communicating with them or AI was ghost-writing instead of him.

*“At first, there were a lot of fans who misunderstood, thinking that ‘Is this (Private Message) also AI?’ [P7’s idol name] explained it wasn’t. He posted a picture of himself to clarify that the messages were sent by [P7’s idol name]. Not just us; I think other fandoms might have initially misunderstood it too. When I received a private message after listening to the call, I wondered, ‘Isn’t this message also sent by the robot?’” (P7, F)*

As the Private Message service was served to paid subscribers alongside with AI Private Call, such confusions in that service were more obvious than other media channels, even though idols were sending by themselves. Text formats contained less identifying information about the senders, so they speculated that text-based social media channels (e.g., X, Instagram) could easily deceive them like AI Private Call.

**Message content.** Participants (P7, P8, P11) stated that AI Private Call users had once been accused on one of the selectable scenario options such as ‘romance’ on social media due to its inappropriateness. They viewed the scenario as inappropriate, violating the norm of the idol-fan relationship, since some fans could overreact to the pseudo-romantic relationship.

*“I am not sure, but from I’ve heard from those around me, there was a scenario that involved a pseudo-relationship. The voice message treated them as if they are in a real romantic relationship, leading to some negativity in the experience of the people involved.” (P7, F)*

As P7 stated, some fans shared the excerpt of the script on social media, which went viral within fandom about its suitability to use. P9 and P11 also noted that some fans found certain script lines potentially inappropriate from a gender perspective (e.g., “Thank you for choosing me”, “Why didn’t you get my call?”). Eventually, the scenario and the script were removed and revised by the company.

**5.2.4 [Intensification] Examine Idol’s Involvement in Content Creation.** Nine participants wanted to know the detailed process of how the company created simulated voice messages with idols (P1, P2, P3, P5, P8, P9, P10, P11, P14). Some of them indicated that they watched promotion videos of the service for how their idols were involved to understand the process of voice creation. They attempted to approximate how they worked to create synthetic voices, rather than deeply engaging in listening to what the message is about (P1, P5, P7, P8, P10). This aligns with fans’ behavior associated with the ‘intensification’ stage, who seek deeper intimacy with celebrities through affirming the authenticity of celebrities on media content [92].

Fans generally engage with gossip and seek relational information to verify the authenticity of celebrities’ behaviors exhibited on social media, including personal disclosures and emotional expressions [41, 56]. With simulated voice messages, on the other hand, participants were motivated to verify the authenticity of the simulated voice message (if their idols are actually committed), as it is obviously artificial-generated content and purportedly endorsed by their idols.

**Voice content.** While the output they listened to was synthetic voices which they could not admit as idols’ voices, they (P5, P7, P8, P10) attempted to approximate idols’ effort in creating these simulated voices. They considered how idols would have behaved to produce these simulated voice messages, based on their idols’ typical behaviors shared on social media content. P7 and P10 stated that their idols must have put a lot of effort into recording their voices when the company created voice models. This echoes prior research on fans’ behaviors, where fans deduce celebrities’ personalities, thoughts, and feelings from their high level of engagement [5, 92].



*"[After listening to simulated voice messages] I just thought, 'He must have worked hard'. In any way, he must have recorded the sound. I thought that he must have worked hardly to record a sound that would be enough for AI to come up with this kind of sentence. I was like, 'It's hard to earn money.'" (P10, F)*

Approximating idols' commitment led them to reluctantly favorably accept AI Private Call even though the experience was not satisfying. This is because they supported their idols' growth and wished them get rewarded, rather than get criticized. P7 stated that she watched a video in which her idol introduced the service, and it led her to understand the extent of his efforts. She expressed dissatisfaction with the experience but still favored AI Private Call, believing her idol enjoyed creating his synthetic voice.

**Message content.** Message content was related to the script of simulated voice messages. Participants wondered if their idols also knew the scripts (P11) or if the scripts of simulated voice messages had been written by their idols (P1, P5, P14). They stated that if they had known the scripts of simulated voice messages were written by their idols, they would have concentrated on listening more profoundly.

*"Let's say he (P1's idol) wrote the entire script. Then, I would have listened to it even if they're (simulated voice messages) awkward or they're not that realistic. Because it is what he's trying to tell me. But I was a little reluctant to listen to the computer's voice because that fact was unclear. If he had written the script, that means he put effort into it. I might have thought, 'What they say is what he says to root for me,' However, I think I was avoiding it more because it wasn't clear how far the technologies had been applied." (P1, Chose not to disclose)*

The entire process of how simulated voice messages made was unclear in this service, except for only providing a promotion video provided by the company. In addition, this is related to K-Pop idol culture, awareness on an idol's involvement makes products or services more reliable and meaningful for fans. For example, voice-recorded content holds greater value than simulated voice messages, because voice-recorded content implies that idols invested more time using their voices, while synthetic media is created effortlessly.

*"It is like the photo cards in the album. There's a big difference in price between the selfie photo cards and the photo cards that others took. They are different in that one is a picture taken by idols on their own and another is taken by someone else. A recorded voice means the member recorded it and made it with his real voice. But the AI voice isn't his voice. That's why I was not satisfied, feeling such a sense of separation from it (simulated voice message)." (P1, Does not wish to answer)*

From the another aspect, this theme is related to participants' concerns about possible violation of idols' personal rights caused by the misuse of simulated voice messages (P1, P2, P7, P8, P10, P11, P13). While such a service using idols' synthetic voices can be created through an official contract with idols, they perceived that it would be unavoidable for idols to create content they might not want because of a system and stakeholders around them. Participants had recognized well about the nature of how the K-Pop idol had been systematically managed, based on their extensive observations during long-term fan activities. For example, management agencies usually manage idols' publicity rights, which authorize commercial use of personal personas in promoting commercial products [34]. P8, P10, and P13 wondered whether the idols agreed to generate things that they had not said personally for the service; *"I wondered whether the idol would be okay with using their own voice to deliver messages they didn't intend to say."*(P8) P10 pointed out that the idols'

power to demand their rights varies depending on their popularity, perceiving that newly debuted idols may not be able to claim their rights in creating simulated voice messages.

Moreover, P7 and P11 were anxious that some malicious fans could potentially utilize simulated voices to generate fan-made content which can influence idols' reputations and popularity in real-life as fake information. On the other hand, some fans may be obsessed with simulated interpersonal behaviors, not distinguishing it from the actual behaviors of idols (P1, P3). This could require idols to adhere to the simulations. They argued that the use of synthetic voices for adjustable interactive messages should involve careful consideration on potential impact on idols' lives from their perspectives.

*"I feel anxious that some people may record it (AI Private Call voice message) deliberately to claim that 'I am his ex-girlfriend, and this is the list (of recording files) (...)' . Since there are some people even spreading fake captures of personal messages (which are synthetically edited) these days, I am very concerned about the possibility of rumors produced by the AI voice." (P7, F)*

**5.2.5 [Intensification] Seek to Enhance Feeling of Social Presence.** Fans in the 'intensification' stage enjoy illusory reciprocal interactions with media figures, which lead them to seek further ways to experience close relationships such as following celebrities' social media accounts [92]. Being exposed to celebrities' information through social media frequently allows fans to feel that they are 'being together' and 'being connected' [39]. Although in the minority, several behaviors as ways to enhance the feeling that idols are socially present were observed from some participants. Interestingly, some participants sometimes engaged in creating a sense of intimacy through the construction of pseudo situations by adjusting content as they preferred.

**Voice content.** Only two participants (P5, P7) stated that they listened to idols' voices on other media to reduce a sense of deception they had experienced from simulated voice messages, after they used the service. Their behavior was a way to fulfill their desire to interact with idols, similar to how simulated voice messages had engaged with them. For example, they watched live streaming recordings or listened to songs, where they could hear their idols' real voices. People feel a sense of social presence and credibility more when they listen to humans' voices than machines' voices [88].

*"I like it (AI Private Call) because it's [P7's idol's name], but it's not the real [P7's idol's name], so after listening to it for a bit, it made me want to see him more, making me feel emptier. I'd rather hear the real [P7's idol's name] voice. Then I just listened to music or watched V-Live (K-Pop live streaming platform) one more time, wishing that the real [P7's idol's name] would come..." (P7, F)*

Similarly, P2 stated that her affection for her idol increased after listening to AI Private Call, as she felt more authentic in the presence of idol on other media channels: "The perception of the idol, and therefore the perception of [P2's idol name] himself, has never changed (with this listening experience), because I know that the two entities have nothing to do with each other, but I think I like him more after Private Call, thinking that 'he is real'" (P2, F)

**Message content.** Some participants (P1, P2, P3, P4, P5, P6, P8, P15) sought ways to enhance a sense of closeness with idols in using AI Private Call. Adjusting simulated voice messages with interface options was a way for them to increase idol-likeness of simulated voice messages (Table 3). In CMC platforms like Twitter, users engage with celebrities by merely following and responding to their activities [72], but in AI Private Call, they proactively engaged with celebrities by tailoring content to shape desired situations through simulated voice messages.

P1, P3 and P4 stated they selected names and speech levels based on how their idols usually interact with them via media channels, or based on how they interact with idols. "There is a name

Options in interface	How participants adjusted AI Private Call voice messages	Purposes
Speech Styles	Choose 'informal' or 'formal' option that matches the idol's actual speech level	To feel like the idol's AI voice as a real person.
	Choose 'informal' or 'formal' option considering the Korean relationship hierarchy between participants and an idol.	To feel like a close relationship with an idol.
	Choose 'informal' or 'formal' option which is the opposite speech level of an idol.	To hear how an idol talks in a way that they don't usually talk to fans.
Scenarios	Choose 'informal' or 'formal' option depending on participants' preference	A comfortable sound for participants.
	Choose scenarios (birthday, morning call, and workout) in specific situations (e.g., work outs).	To increase participants' self-efficacy.
	Choose scenarios (birthday, morning call, and workout) in which the voice is pronounced naturally.	To feel like the idol's AI voice as a real person.
Name	Designated as participants' first name with the suffix (ya), 'Noo-na'* or 'Oppa'* considering the Korean relationship hierarchy between participants and an idol.	To feel like a close relationship with an idol.
	Designated as a word that voice pronounced naturally	To feel like the idol's AI voice as a real person.

Table 3. The case about how the participants adjust the simulated voice message with options of AI Private Call. \*Korean suffixes that are typically added to the first name [83].

that [P3's Idol group name] tends to call their fans. Since the name of [P3's Idol group name]'s fandom is 'Fearless', so I used the word [Fearless] as a name." (P3, M) In particular, he stated that he adjusted a speech level differently in the options, depending on each idol's perceived personality. P4 aligned his use of AI Private Call with his presence in fandom communities, such as setting his name as same as the one he used in online community. This was also because they wanted to listen to simulated voice messages as if their idols were talking to them in natural, as P4 clarified. Fans desire their celebrities to be consistent across media channels [98].

On the contrary, some participants (P2, P3, P5, P6, P8) stated that they chose options in such a way that they could immerse themselves in personal situations. They generally preferred choosing 'Informal speech level' where simulated voice messages were delivered in an informal tone. Also, considering how they would be called if they are in personal relationships, they changed the name by adding the suffix to their first name (a/ya: 아/야, e.g., Jennie-ya) or Noo-na and Oppa (e.g., Jennie Noo-na: 누나), which are usually used among people who are close to others in Korea [83].

*"I tended to select 'informal', because the idols rarely talk down in informal ways. When they communicate with the fans, they use honorifics unless the fans ask them to speak informally... Even if we don't use honorifics, [P2's idol name] always uses honorifics for us. [...] While he typically uses honorifics a lot, I feel a bit more familiar when he uses informal speech. Having a personal conversation with an idol suggests that I*

want to feel a personal connection between us. That's why I chose the speak option as 'informal.'" (P2, F)

In addition, a few participants (P2, P4, P5), who were the most active users, stated that they deliberately selected scenarios that corresponded to their current situations, where they could immerse themselves in the simulated voice messages. For example, P4 chose the 'Cheer up' scenario when they needed an energy boost, such as during weight training. While idols could cheer on their fans via live-streaming or their social media channels, it is challenging for individuals to receive personalized encouragement like AI Private Call whenever they wanted. They selected various scenarios, such as 'Birthday' for celebration and 'Morning call' to wake up early (P2). By adjusting these call options to their situations, they stated that they felt energized, as if they were being cheered on by their idols (P4, P5). Their reactions align with the previous research showing that individuals who had a high level of parasocial relationship with a doctor character on television felt an increase in self-efficacy [75]. There was a case where a participant chose options which they could feel less awkwardness than other options (P15).

"Especially when I received similar answers during my workout, I felt genuinely thrilled. They truly cheered me on. In those moments, I realized that I needed to put in more effort and work harder." (P4, M)

5.2.6 *[Integration] Sharing Content Experience with Fandom.* Sharing experiences with like-minded individuals is a common behavior for fans in the 'integration' stage [92]. Fans gather and collaborate with each other to intensify their intimacy with celebrities, such as socializing with like-minded individuals, collecting memorabilia, and publishing fan-made content [81, 92]. Similarly, participants also wanted to enjoy the content with other fans, which led them to encourage others to join them. While it was a minority, some participants (P2, P4) who had positive impressions from the service attempted to share their thoughts within their fandom communities. One participant (P2) shared her experience after recording the best part of the message she thought in her private chat with other fan friends, expecting that others would resonate with and enjoy the service as she did.

*"I've shared it (AI Private Call) a few times. But many people around me were negative about the service, like 'This is some kind of deepfake', 'It's a bit uncomfortable', 'Why does it have to use an idol's voice?', and 'That's fake'. Whenever I tried to share my experience, reactions were not that positive, so I began not to share them further. If there was a good conversation to share, I would have shown it (simulated voice message) to people, like 'hey, try it'..."* (P2, F)

The positive impression prompted P2 to share the content with other members, but she mentioned her hesitancy because of others' anticipated pessimistic reactions. As opposed to P2, P1, and P8 stated that they playfully mocked those who might be interested in listening to 'fake' idol voices with other fans.

Fans would like to create, curate, and share playful content related to their idols with other fans to feel fun together [78]. For example, fans usually capture interesting social media posts of idols, then they share them with other fans to enjoy together. P4 stated that he had been using a community board (DC inside) where his fandom primarily communicated. He wrote about how to use AI Private Call effectively to promote the positive aspects of the service to other fan members. However, as P2 stated above, user experience of AI Private Call was not notable to share for most participants, so it was rare to find participants who reached this stage. P8 used Private Message as an example, noting that its content is easily shareable due to its traits that facilitate sense-making for other fans.

*"[Private Message] is simple to share with others. The biggest difference is that this Private Call is a fake, and aside from that, the script of the Private Call can be roughly predictive.*

*But another service (Private Message) plainly presents the characteristics of this idol. You can see their (idols) personality and characteristics so well. I think that's why the service is more fun, [...] There aren't many responses in Private Call. When I first started this service, I could only share it to say like, "Try to listen how similar it is." (P8. F)*

## 6 Discussion

We examined how fans perceive simulated voice messages that imitate idols (RQ1) and how they cognitively and behaviorally respond to simulated voice messages in conjunction with parasocial relationships (PSRs) establishment (RQ2). While this study provides a glimpse into fans' reaction to a K-Pop idol fanbase service, we believe that the findings offer valuable insights necessary for the impact of synthetic media involving individuals who hold influential connections with the audience. First, we examine how parasocial relationships can be formed through simulated voices. Next, we explore the impact of fandoms, such as K-Pop, on individual fan experiences. We also address the potential risks synthetic voices pose to celebrities. Finally, we propose design implications of synthetic media fans can engage with.

### 6.1 Parasocial Relationship Establishment mediated by Simulated Voice

Our findings indicate that fans engage with simulated voice messages mimicking celebrities, thereby developing parasocial relationships. In particular, from content that simulates 'communication behavior', which performs imitated a celebrity's interpersonal interactions, fans can feel parasocial connections (Finding 5.1.2). Previous research indicates that a speaker's engagement in sharing personal details can influence parasocial relationships (PSRs) [80]. Our study extends these findings, showing that synthetic celebrity voices exhibiting similar behaviors can also evoke parasocial attractions. However, the nature of synthetic voices, which people might often perceive as fake (Finding 5.1.1) and social contexts of fans that demand authenticity from their idols, can inevitably diminish the quality of listening experience. Moreover, how voice messages perform is interpreted based on social contexts surrounding fans, such as how the idols interact with them or how their communities accept such technologies. While Chan et al. [10] demonstrated that Voice User Interface using familiar voices were perceived as more engaging and persuasive, this result may not always be consistent in the fan culture context.

Utilizing a theoretical framework that shows how audiences can develop parasocial relationships through media content [92], we show that fans can engage with simulated voice messages via series of stages to establish parasocial relationships (Figure 2). The engagement process unfolds dynamically, starting from the initiation stage. In our findings, only a few satisfied participants advanced to the following stage, and fans can sometimes show behaviors associated with multiple stages concurrently. Our work helps to understand how fans would perceive both the voice and content communicated by synthetic celebrity voices, and what information would be utilized to engage them.

In comparison with mass media (e.g., TV, radio) and computer-mediated communication (CMC) media (e.g., online chats) which assume that celebrities would be authentically present, we note that synthetic celebrity voices can foster distinctive behaviors for parasocial interactions, based on the findings. First, our findings showed that participants examined technical implementation of synthetic voice content (How the content includes idols' characteristics well)(Finding 5.2.1, 5.2.2), which has not been observed in the other media. In general media, audiences gather information about attractive media figures such as understanding the 'concept of the character' [90], rather than heuristically examining the structure. This distinct behavior may be attributed to the lack of authenticity in synthetic voices when mimicking celebrities [26]. Individuals engage with interactions in communication, based on the authenticity they evaluated [46]. Knowing the content

was fake, participants may have adjusted their level of engagement based on their evaluations. This propensity has been also deemed crucial for other virtual entities. To ensure authenticity in their presence, virtual entities (e.g., virtual YouTubers) often maintain consistent identities and communication styles across media channels [13, 54].

Second, the progression of PSR stages through synthetic celebrity voice content can vary, regardless of how long people maintain their fan-ship. For example, P2 shared her experiences with other fans, while P7 did not, even though both had been fans for a long period of time. Cambre and Colnago et al. [9] demonstrated that the perception of synthetic voices' suitability for a specific format is subjective. As they demonstrated, participants' general voice preferences and how much they desire the concept might have influenced satisfaction with simulated voice messages. However, interestingly, certain high engagement level towards idols may make the listener attempt to overcome the negative experiences. P7 expressed negative views about the probe, while also grappling with these views due to their attachment to idols (Finding 5.2.4). Understanding how prior impression on a represented person influences the ability to overcome negative and eerie media experiences would be an interesting question to understand how to effectively communicate synthetic voice content.

Lastly, our findings show that simulating celebrities may augment imagined interactions that fans engage with celebrities (Finding 5.2.5). We focused solely on simulated voice messages, where the user can only listen passively. However, if the use case expands to include generative AI that can verbally interact with users as intended, it may deepen user engagement by allowing them to customize their experiences, as demonstrated by some participants (Table 3). This can be a unique characteristic of synthetic media content, distinct from the general mass media where parasocial interactions are often referred to as 'one-sided interactions,' with audiences passively consuming content [26]. Even though CMC media provides direct access to celebrities, the relationships still remain mass-personal [69], with celebrities retaining the initiative to share information. As participants agreed that simulating celebrities would be able to diversify entertainment options for fans (Finding 5.1.2), we anticipate that interaction designs, such as enabling interactive conversations, would create more opportunities for some media figures to increase fan engagement. However, as the Finding 5.2.5 noted, audience's perception of social presence from simulated celebrity content would not match the celebrities' intended self-presentation. This could foster problems, such as people believing the simulation is real or expecting celebrities to perform like the simulation. For example, fan-made songs using AI-simulated voices of artists often incorporate elements from the artists' personal lives or various genres, which may not align with the artists' preferences.

Our research highlights that using simulated celebrity voices can create diverse interactions of individual fans, which can potentially influence target speakers and their communities. Technology designers should comprehend potential impacts that design might have on parasocial relationships and how to seamlessly incorporate designed components into the users' social contexts.

## 6.2 Fandoms that Contribute to Parasocial Relationship Development

Our findings show that fandom plays a significant role across the PSR development. Previous research stressed how communal experiences, including exchanging feedback, promoting activism, and sharing playful interactions, are formulated in fandom communities [2, 11, 22, 78]. Our research concurs with prior studies, showing that individual fan experiences are influenced by both personal interactions with content and engagement within communities. In particular, fan communities serve as resources for fans to obtain reliable information about idols (Finding 5.2.3). This is consistent with the work of Lee et al. [47] and King et al. [41], which suggest that fandoms verify the credibility of online information, including idols' emotional expressions and misinformation. Interestingly, our findings show that the type of information individuals rely on can vary depending on their

community size and the sense of belonging as a community member. For instance, most participants used social media platforms like X to gain a broad understanding of fan opinions on the probe, whereas they obtained the probe's reliability from exclusive fandom community boards or smaller friend networks (Findings 5.2.3). They also used celebrity accounts to directly address and dispel rumors about the probe. Sharing experiences were also attempted in the closer fan network (Findings 5.2.6). How users perceive reliability of information on social media has been studied focusing on healthcare and politic communication [29, 91]. Future research of how individuals utilize associated communities as reliable resources and how they want to contribute to the communities would be conducive to understand information seeking behaviors around fandom services.

### 6.3 Different Personalities of K-Pop Fans that Affect Experiences

All of our findings are derived from K-Pop fans, who often foster strong relationships with their idols and fellow fans through active social media activities [55, 78]. Additionally, the genuine emotional connection expressed by celebrities is crucial in motivating fans to work for them (e.g., voting for music awards) [41, 89]. These characteristics could be related to our participants' behaviors and cognition, differing from other fan cultures such as sports. For example, examining information in online channels (Finding 5.2.3) can be more active in K-Pop fan cultures, as they have substantial online communities where they can gather information. Also, we expect behaviors or cognition reflecting various cultural norms to be demonstrated, particularly depending on how authenticity is shown. For example, idols can demonstrate their authenticity by showcasing the effort they put into creating synthetic media content, which reflects their general emotional expressions towards fans [41] (Finding 5.2.4). On the other hand, TikTok influencers who often display socially accepted emotional 'rawness' to strategically convey authenticity because they usually communicate with their viewers through short-videos recommended by algorithms [6]. In communities, fans can also be categorized by their personalities, such as followers and coworkers [36]. Thus, such communities may expect various ways to demonstrate authenticity.

We also anticipated that the public, or non-fans might have different perspectives on simulated voice messages and understanding on a sense of authenticity of celebrities. They tend to consume content passively and casually without forming attachments to celebrities [35]. Hence, non-fans might remain indifferent to the gap between synthetic voices and real voices, as they are not engaged with celebrities. More seriously, they may be deceived easily by AI-generated content of celebrities [70]. Moreover, they are less likely to recognize ethical problems, compared to fans who take care of protection of idols from potential misuse of AI technologies.

### 6.4 Potential Risks Posed by Simulated Voice Messages

In line with recent concerns about the use of AI technologies to simulate media figures [30, 94], our findings introduce potential problems voice cloning technologies could pose after post-deployment (Finding 5.2.4). In real world, misinformation disseminated by the press and anti-fans has harmed celebrities, with some fans occasionally stalking their personal lives [57, 85]. In addition, celebrities sometimes face inappropriate interactions with fans who are deeply obsessed [31]. These issues could become severe if synthetic celebrity voices become realistic and their sub-features allow users to deeply engage with the simulations. Additionally, the use of synthetic voices could make conspiracy theories that confuse intended communication partners, leading them to mistakenly suspect real people as AI on social media (Finding 5.2.3). As Hutiri et al. [33] noted, extensive harms caused by speech generators can occur to a person who is a subject of and interacts with over a long time.

One approach could involve enhancing competencies of fandoms to recognize potential problems that may be caused by AI-based celebrity content. Long and Magerko [53] suggested that AI literacy

should be designed considering learners' identity and values. Platforms or content providers could collaborate with fandoms to address problems and assess their severity from the fans' perspectives. While fandoms have been previously considered minor social groups, they have been critically interpreted media content, sometimes reworking them for their cultural creations [35]. We believe involving fans in the voice design process can contextualize solutions that researchers currently suggest to protect target speakers from voice cloning attacks [52]. This approach provides insights into how the designed components fit their cultural knowledge.

## 6.5 Design Implications for Producing Synthetic Media Emulating Celebrities

Participants suggested several design implications for synthetic celebrity voices. Also, their reflections on how they used the service have implications for what they expected from simulated messages to engage with celebrities. We present these implications that help the design process of celebrities' AI-simulated voices to be more authentic and engaging.

*6.5.1 Channeling the Celebrity's Media personalities.* Participants expected synthetic voices to have idols' usual speech styles exhibited on social media, and the message content to contain idol's general habits (Finding 5.2.1). Across various media channels, celebrities maintain their own desired personalities, channels communicating with audiences, and modes of appealing to audiences [15, 93]. Additionally, various media sources and communities can shape these specific personalities and traits of them. Thus, synthetic voices should be selected to align with perceived media personalities and their roles in shaping these personalities. For example, as P15 noted, the synthetic voice content could be scheduled to align with celebrities' publicly announced availability.

*6.5.2 Involving Celebrities to be Situated as Trustworthy Content.* Participants expected their idols to be the most involved in content creation and to be aware of how fans experienced (Finding 5.2.4). Even if the output is generated by AI, celebrities could participate in the design process from the scratch and keep communicating with their fans about what they committed in design process, addressing perceived problems. This would increase the reliability of the content, reducing confusion in recognizing celebrities on other social media platforms (Finding 5.2.3). At the same time, how celebrities perceive their simulated content could allow fans to engage with the content to empathize with them. For example, when an idol shared his experience of AI Private Call, laughing at how his AI voice sounded, his fans also tried to use the service to have fun together [104]. Likewise, the role of idols to integrate the content with fan communities may be conducive in strengthening authenticity of the content.

*6.5.3 Regulating the Usage of the Celebrity's Data and Communicating Its Transparency.* Participants had concerns about the potential violation of their idols' personal boundaries, which resulted in a distracting user experience (Finding 5.2.4). Celebrities themselves would care for their personality rights the most, such as actors striking against using their data to create simulated avatars [43], and how celebrities' data is cared can also influence fans in engaging with the content as they do not want to harm their favorites due to their entertainment. Nevertheless, unintended attack such as NSFW (Not safe for work) could occur by some fans. Practitioners should create appropriate regulations to internally and externally safeguard celebrities, being transparent about the model production process - from data collection to data training for the AI models. Existing approaches from AI-infused systems, such as model cards [59], could guide what information should be disclosed. Content providers should also provide comprehensive explanations of protective measures for celebrities from malicious users who may generate synthetic media (e.g., pornography) of celebrities to damage celebrities' reputations. Additionally, as AI technologies become more autonomous, they might produce content that celebrities did not intend. Therefore, practitioners should consider the



permissible scope of such generation. They could also give the authorization to fans to opt-out of the use of synthetic media in services for those who do not want to use their data.

*6.5.4 Providing Controllability to Adjust Content.* Some participants mentioned they felt more immersed by adjusting voice messages to suit interpersonal relationships (Finding 5.2.5). As shown in Table 3, they selected different options to customize content, focusing on either realism or fantasy based on their preferences. Thus, technology designers could consider if users are allowed to be given autonomy to enhance fan engagement, such as designing conversational AI with celebrity voices or interactive features that users can change styles of voices. This would enable users to engage in dynamic dialogues and form emotional bonds with them [28]. However, as a trade-off, this user-side control might also generate unintended content, potentially allowing the users to generate content harming the target speakers. Thus, this implication should be considered alongside other design implications, especially those in Section 6.5.3 to reduce its risks towards celebrities.

*6.5.5 Structuring Cultural Values Necessitated by Fan Communities.* Our findings show that fans predominantly seek their fandoms' opinions when exploring how the probe could enrich their communities (Finding 5.2.3, 5.2.6). While the probe was limited to individual experiences with personalized simulated voice messages, participants wanted to connect with other fans and contribute to their communities by sharing content. Such behaviors could also be observed in other fan communities such as sharing friendship bracelets between Taylor Swift fans [101]. This is because fans build a participatory environment in which all fans playfully interact, centered around their favorites [78], and they desire to socially interact with others having the same cultural identities [35, 84]. Therefore, technology designers should understand culturally acceptable values that fans enjoy and how to encourage playful interaction within fan communities. Depending on fandoms' characteristics, entertainment may be performed differently. For example, 'Learn! KOREAN with TinyTAN' [100] allows foreign K-Pop fans to learn Korean with BTS members' AI generated voices, which encourages fans to communicate with other fans and understand what the members say.

## 6.6 Limitation

We acknowledge several limitations around our research. First, our results may be biased due to an imbalance in participant demographics. We posted a call for participants across diverse fandom communities of idols served by AI Private Call to minimize the potential bias arising from differences between each fandom community. However, the application rates varied depending on the characteristics of fandom communities. Consequently, seven out of our fifteen participants were from the same fandom (Brave Girls). Second, we did not interview fans who were not highly involved or engaged within the fandom. AI Private Call was only available to paid subscribers of the Universe app, meaning that all our interviewees were willing to invest in a subscription with substantial level of fanships. As such, all of our participants were relatively highly engaged with the idols and the fandom. However, there may be fans who identify as such but have a lower level of engagement due to economic reasons, different priorities, or other factors. Third, the generalizability of the findings may be limited due to the nature of this study. The fans we interviewed used the same native language as the idols, and the fandom culture was in the context of K-pop idols. As we noted in the discussion section (Section 6.3), K-Pop fandoms have their own uniqueness in socializing with others via social media, how the listeners process simulated voice messages could differ depending on how they obtained information or perceive their media figures. Future research should extend to fans who use different languages and fandoms beyond K-pop idols. Lastly, we chose the 'retrospective interview' method where interviewees walk-through their

past experiences. Our goal was to gain valuable insights from their usage periods to answer RQs. In this process, interviewees might have been biased in reflecting on their experiences or limited in refining them.

## 6.7 Conclusion

In this work, we explored how simulated voice messages from idols influence parasocial relationships by understanding fans' perceptions, cognitions, and behaviors related to these simulated messages. We conducted interviews with AI Private Call users, inquiring about their experiences using the service. Utilizing a theoretical model outlining four stages (initiation, experimentation, intensification, and integration) of parasocial relationship development, our results underscore that fans show a series of actions to nurture their PSRs through simulated voice messages. We present the perceptual process detailing how fans process the experience of simulated voice messages and discuss the considerations necessary in the era of synthetic media.

## Acknowledgments

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. RS-2024-00406715) and by Institute of Information & Communications Technology Planning & Evaluation (IITP) grant funded by the Korea government (MSIT) (No. RS-2024-00443251, Accurate and Safe Multimodal, Multilingual Personalized AI Tutors).

We sincerely appreciate reviewers who helped enhance our paper and participants who shared their valuable experiences for our studies. We also thank all members of the KAIST Interaction Lab (KIXLAB) and the Cornell Communication and Collaborative Technologies Lab for their support.

## References

- [1] Michael E Agee. 2014. *I listen to you every day: Parasocial relationships and self-disclosure in Christian radio*. The University of Texas at Arlington.
- [2] Alyse Marie Allred and Colin M Gray. 2021. "Be Gay, Do Crimes": The Co-Production and Activist Potential of Contemporary Fanzines. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW2 (2021), 1–35.
- [3] Matthew P Aylett, Leigh Clark, Benjamin R Cowan, and Ilaria Torre. 2021. Building and designing expressive speech synthesis. In *The Handbook on Socially Interactive Agents: 20 years of Research on Embodied Conversational Agents, Intelligent Virtual Agents, and Social Robotics Volume 1: Methods, Behavior, Cognition*. 173–212.
- [4] Matthew P Aylett, Benjamin R Cowan, and Leigh Clark. 2019. Siri, echo and performance: You have to suffer darling. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–10.
- [5] Murray R Barrick, Gregory K Patton, and Shanna N Haugland. 2000. Accuracy of interviewer judgments of job applicant personality traits. *Personnel Psychology* 53, 4 (2000), 925–951.
- [6] Kristen Barta and Nazanin Andalibi. 2021. Constructing authenticity on TikTok: Social norms and social support on the "Fun" Platform. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW2 (2021), 1–29.
- [7] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative research in psychology* 3, 2 (2006), 77.
- [8] Nancy H Brinson, Laura L Lemon, Coral Bender, and Annika Fetzer Graham. 2023. Consumer response to podcast advertising: the interactive role of persuasion knowledge and parasocial relationships. *Journal of Consumer Marketing* 40, 7 (2023), 971–982.
- [9] Julia Cambre, Jessica Colnago, Jim Maddock, Janice Tsai, and Jofish Kaye. 2020. Choice of voices: A large-scale evaluation of text-to-speech voice quality for long-form content. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–13.
- [10] Sam WT Chan, Tamil Selvan Gunasekaran, Yun Suen Pai, Haimo Zhang, and Suranga Nanayakkara. 2021. KinVoices: Using voices of friends and family in voice interfaces. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW2 (2021), 1–25.
- [11] Ruijia Cheng and Jenna Frens. 2022. Feedback exchange and online affinity: A case study of online fanfiction writers. *Proceedings of the ACM on Human-Computer Interaction* 6, CSCW2 (2022), 1–29.

- [12] Chau-kiu Cheung and Xiao Dong Yue. 2012. Idol worship as compensation for parental absence. *International Journal of Adolescence and Youth* 17, 1 (2012), 35–46.
- [13] Abhinav Choudhry, Jinda Han, Xiaoyu Xu, and Yun Huang. 2022. "I Felt a Little Crazy Following a'Doll'" Investigating Real Influence of Virtual Influencers on Their Followers. *Proceedings of the ACM on human-computer interaction* 6, GROUP (2022), 1–28.
- [14] Siyoung Chung and Hichang Cho. 2017. Fostering parasocial relationships with celebrities on social media: Implications for celebrity endorsement. *Psychology & Marketing* 34, 4 (2017), 481–495.
- [15] Cinzia Colapinto and Eleonora Benecchi. 2014. The presentation of celebrity personas in everyday twittering: Managing online reputations throughout a communication crisis. *Media, Culture & Society* 36, 2 (2014), 219–233.
- [16] Robert Dale. 2022. The voice synthesis business: 2022 update. *Natural language engineering* 28, 3 (2022), 401–408.
- [17] DC Inside. 2024. DC Inside.com. <https://www.dcsinside.com/> Accessed: 2023-07-20.
- [18] DeepL. 2024. DeepL Translator. <https://www.deepl.com/en/translator> Accessed: 2023-09-20.
- [19] Tiffany D Do, Ryan P McMahan, and Pamela J Wisniewski. 2022. A new uncanny valley? The effects of speech fidelity and human listener gender on social perceptions of a virtual-human speaker. In *Proceedings of the 2022 CHI conference on human factors in computing systems*. 1–11.
- [20] Siobhan Doyle. 2021. License your voice. *Engineering & Technology* 16, 8 (2021), 79–84.
- [21] Stuart Dredge. 2020. New K-pop group Aespa includes both human and virtual members. <https://musically.com/2020/10/29/new-k-pop-group-aespa-includes-both-human-and-virtual-members/> Accessed: 2023-09-10.
- [22] Brianna Dym, Cecilia Aragon, Julia Bullard, Ruby Davis, and Casey Fiesler. 2018. Online fandom: Boldly going where few CSCW researchers have gone before. In *Companion of the 2018 ACM Conference on Computer Supported Cooperative Work and Social Computing*. 121–124.
- [23] Jon Fingas. 2020. Björk and Microsoft use AI to create music that changes with the sky. <https://www.engadget.com/2020-01-17-bjork-and-microsoft-ai-sky-music.html> Accessed: 2023-09-10.
- [24] Sam Ford. 2014. Fan studies: Grappling with an 'Undisciplined' discipline. *Journal of Fandom Studies* 2, 1 (2014), 53–71.
- [25] Kathleen M Galotti. 2017. *Cognitive psychology in and out of the laboratory*. Sage Publications.
- [26] David C Giles. 2002. Parasocial interaction: A review of the literature and a model for future research. *Media psychology* 4, 3 (2002), 279–305.
- [27] Leo A Goodman. 1961. Snowball sampling. *The annals of mathematical statistics* (1961), 148–170.
- [28] Juhye Ha, Hyeon Jeon, Daeun Han, Jinwook Seo, and Changhoon Oh. 2024. CloChat: Understanding How People Customize, Interact, and Experience Personas in Large Language Models. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. 1–24.
- [29] Daniel Halpern and Jennifer Gibbs. 2013. Social media as a catalyst for online deliberation? Exploring the affordances of Facebook and YouTube for political expression. *Computers in human behavior* 29, 3 (2013), 1159–1168.
- [30] Robert Hart. 2024. Here's Why Other Celebrities Could Face Problems With AI Voice Cloning—Not Just Scarlett Johansson. <https://www.forbes.com/sites/roberthart/2024/05/24/heres-why-other-celebrities-could-face-problems-with-ai-voice-cloning-not-just-scarlett-johansson/> Accessed: 2024-07-14.
- [31] Jens Hoffmann and Lorraine Sheridan. 2008. Celebrities as victims of stalking. *Stalking, threatening, and attacking public figures: A psychological and behavioral analysis* (2008), 195–213.
- [32] Emma Holmes, Ysabel Domingo, and Ingrid S Johnsrude. 2018. Familiar voices are more intelligible, even if they are not recognized as familiar. *Psychological science* 29, 10 (2018), 1575–1583.
- [33] Wiebke Hutiri, Orestis Papakyriakopoulos, and Alice Xiang. 2024. Not My Voice! A Taxonomy of Ethical and Safety Harms of Speech Generators. In *The 2024 ACM Conference on Fairness, Accountability, and Transparency* (Rio de Janeiro Brazil, 2024-06-03). ACM, 359–376. doi:10.1145/3630106.3658911
- [34] Cornell Legal Information Institute. 2024. publicity. <https://www.law.cornell.edu/wex/publicity> Accessed: 2023-09-10.
- [35] Henry Jenkins. 2012. *Textual poachers: Television fans and participatory culture*. Routledge.
- [36] Soyeon Jin, Saerom Lee, and Hakyeon Lee. 2024. How do fandom types differ? A taxonomy of K-pop fandom with network embedding. *Entertainment Computing* (2024), 100767.
- [37] Spike Jonze. 2013. Her [Film]. Annapurna Pictures.
- [38] Heo Joon. 2021. Launch of UNIVERSE: NCsoft's Advanced Technology Connects Fans and Artists. Accessed: 2023-09-10.
- [39] Jihyun Kim and Hayeon Song. 2016. Celebrity's self-disclosure on Twitter and parasocial relationships: A mediating role of social presence. *Computers in Human Behavior* 62 (2016), 570–577.
- [40] Jieun Kim and Hayeon Song. 2024. My Voice as a Daily Reminder: Self-Voice Alarm for Daily Goal Achievement. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. 1–16.
- [41] Rebecca Chiyoko King-O'Riain. 2021. "They were having so much fun, so genuinely...": K-pop fan online affect and corroborated authenticity. *New media & society* 23, 9 (2021), 2820–2838.

- [42] Christoph Klimmt, Tilo Hartmann, and Holger Schramm. 2013. Parasocial interactions and relationships. In *Psychology of entertainment*. Routledge, 291–313.
- [43] Ken Klippenstein. 2023. As Actors Strike for AI Protections, NETFLIX List 900,000 AI Job. <https://theintercept.com/2023/07/25/strike-hollywood-ai-disney-netflix/> Accessed: 2023-08-31.
- [44] Elly A Konijn and Johan F Hoorn. 2017. Parasocial interaction and beyond: Media personae and affective bonding. *The international encyclopedia of media effects* (2017), 1–15.
- [45] Ropek L. 2023. Grimes launched a platform to help you make AI songs with her voice. here’s what it sounds like. <https://gizmodo.com/grimes-elon-musk-openai-ai-music-elf-tech-1850409972> Accessed: 2023-06-28.
- [46] Eun-Ju Lee. 2020. Authenticity model of (mass-oriented) computer-mediated communication: Conceptual explorations and testable propositions. *Journal of Computer-Mediated Communication* 25, 1 (2020), 60–73.
- [47] Jin Ha Lee, Nicole Santero, Arpita Bhattacharya, Emma May, and Emma S Spiro. 2022. Community-based strategies for combating misinformation: Learning from a popular culture fandom. *Harvard Kennedy School Misinformation Review* (2022).
- [48] Kwan Min Lee and Clifford Nass. 2003. Designing social presence of social actors in human computer interaction. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. 289–296.
- [49] Rachel Esther Lim and So Young Lee. 2023. “You are a virtual influencer!”: Understanding the impact of origin disclosure and emotional narratives on parasocial relationships and virtual influencer credibility. *Computers in Human Behavior* 148 (2023), 107897.
- [50] Jiahui Liu. 2023. Virtual presence, real connections: Exploring the role of parasocial relationships in virtual idol fan community participation. *Global Media and China* (2023), 20594364231222976.
- [51] Zihao Liu, Yan Zhang, and Chenglin Miao. 2023. Protecting Your Voice from Speech Synthesis Attacks. In *Annual Computer Security Applications Conference* (Austin TX USA, 2023-12-04). ACM, 394–408. doi:10.1145/3627106.3627183
- [52] Zihao Liu, Yan Zhang, and Chenglin Miao. 2023. Protecting Your Voice from Speech Synthesis Attacks. In *Proceedings of the 39th Annual Computer Security Applications Conference*. 394–408.
- [53] Duri Long and Brian Magerko. 2020. What is AI literacy? Competencies and design considerations. In *Proceedings of the 2020 CHI conference on human factors in computing systems*. 1–16.
- [54] Zhicong Lu, Chenxinran Shen, Jiannan Li, Hong Shen, and Daniel Wigdor. 2021. More kawaii than a real-person live streamer: understanding how the otaku community engages with and perceives virtual YouTubers. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–14.
- [55] Marlyna Maros, ABDUL BASEK, and FARA NADIA. 2022. Building Online Social Identity and Fandom Activities of K-pop Fans on Twitter. *3L: Language, Linguistics, Literature* 28, 3 (2022).
- [56] Alice Marwick and Danah Boyd. 2011. To see and be seen: Celebrity practice on Twitter. *Convergence* 17, 2 (2011), 139–158.
- [57] Sarah McRae. 2017. Get off my internets: How anti-fans deconstruct lifestyle bloggers’ authenticity work. *Persona Studies* 3, 1 (2017), 13–27.
- [58] Edvinas Meskys, Julija Kalpokiene, Paul Jurcys, and Aidas Liaudanskas. 2020. Regulating deep fakes: legal and ethical considerations. *Journal of Intellectual Property Law & Practice* 15, 1 (2020), 24–31.
- [59] Margaret Mitchell, Simone Wu, Andrew Zaldivar, Parker Barnes, Lucy Vasserman, Ben Hutchinson, Elena Spitzer, Inioluwa Deborah Raji, and Timmit Gebru. 2019. Model cards for model reporting. In *Proceedings of the conference on fairness, accountability, and transparency*. 220–229.
- [60] Margaret Moe. 2023. Podvertising II: “Just like My Best Friend” – Relationships in Host-read Podcast Advertisements. *30, 1* (2023), 362–386. doi:10.1080/19376529.2021.1960840
- [61] Azadeh Naderi and Donghee Yvette Wohn. 2023. User engagement behaviors: Exploring predictors of ARMY’s active and passive media usage behaviors using social cognitive theory. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–7.
- [62] Naver. 2024. Clova Note. <https://clovanote.naver.com/> Accessed: 2023-09-20.
- [63] Naver. 2024. PAPAGO. <https://www.dcinside.com/> Accessed: 2023-07-20.
- [64] Paarth Neekhara, Shehzeen Hussain, Shlomo Dubnov, Farinaz Koushanfar, and Julian McAuley. 2021. Expressive neural voice cloning. In *Asian Conference on Machine Learning*. PMLR, 252–267.
- [65] Shuo Niu, Hugh S Manon, Ava Bartolome, Nguyen Binh Ha, and Keegan Veazey. 2022. Close-up and whispering: An understanding of multimodal and parasocial interactions in YouTube ASMR videos. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 1–18.
- [66] Nurhafizh Noor, Sally Rao Hill, and Indrit Troshani. 2022. Artificial intelligence service agents: role of parasocial relationship. *Journal of Computer Information Systems* 62, 5 (2022), 1009–1023.
- [67] Camille Noufi, Lloyd May, and Jonathan Berger. 2023. The Role of Vocal Persona in Natural and Synthesized Speech. In *2023 IEEE 17th International Conference on Automatic Face and Gesture Recognition (FG)*. IEEE, 1–4.

- [68] Garrett J O'Keefe and Barbara K Sulanowski. 1995. More than just talk: Uses, gratifications, and the telephone. *Journalism & Mass Communication Quarterly* 72, 4 (1995), 922–933.
- [69] Patrick B O'Sullivan and Caleb T Carr. 2018. Masspersonal communication: A model bridging the mass-interpersonal divide. *New media & society* 20, 3 (2018), 1161–1180.
- [70] Reza Arkan Partadiredja, Carlos Entrena Serrano, and Davor Ljubenkov. 2020. AI or human: the socio-ethical implications of AI-generated media content. In *2020 13th CMI Conference on Cybersecurity and Privacy (CMI)-Digital Transformation-Potentials and Challenges (51275)*. IEEE, 1–6.
- [71] Loretta L Pecchioni and Kelby K Halone. 2000. Relational listening II: Form & variation across social and personal relationships. *International Journal of Listening* 14, 1 (2000), 69–93.
- [72] Natalie Pennington, Jeffrey A Hall, and Alex Hutchinson. 2016. To tweet or not to tweet: Explaining fan-celebrity interaction on Twitter. *Interaction* 48, 1 (2016).
- [73] Greig Powers, James P Johnson, and Ginger Killian. 2023. To Tell or Not to Tell: The Effects of Disclosing Deepfake Video on US and Indian Consumers' Purchase Intention. *Journal of Interactive Advertising* (2023), 1–17.
- [74] Alice Preminger and Matthew B Kugler. 2023. The Right of Publicity Can Save Actors from Deepfake Armageddon. *Berkeley Technology Law Journal, Forthcoming, Northwestern Public Law Research Paper* 23-52 (2023).
- [75] Eric E Rasmussen and David R Ewoldsen. 2016. Treatment via television: The relation between watching Dr. Phil and viewers' intentions to seek mental health treatment. *Journal of health communication* 21, 6 (2016), 611–619.
- [76] Tobias Raun. 2018. Capitalizing intimacy: New subcultural forms of micro-celebrity strategies and affective labour on YouTube. *Convergence* 24, 1 (2018), 99–113.
- [77] ResembleAI. 2023. Twitch TTS - custom AI voices for text-to-speech. <https://www.resemble.ai/twitch/> [Online; accessed 17-Sep-2023].
- [78] Kathryn E Ringland, Arpita Bhattacharya, Kevin Weatherwax, Tessa Eagle, and Christine T Wolf. 2022. ARMY's magic shop: Understanding the collaborative construction of playful places in online communities. In *Proceedings of the 2022 CHI conference on human factors in computing systems*. 1–19.
- [79] Johnny Saldaña. 2021. *The coding manual for qualitative researchers*. sage.
- [80] Daniela Schlütz and Imke Hedder. 2022. Aural parasocial relations: Host–listener relationships in podcasts. *Journal of Radio & Audio Media* 29, 2 (2022), 457–474.
- [81] Jesse Weaver Shipley. 2015. Selfie love: Public lives in an era of celebrity pleasure, violence, and social media. *American Anthropologist* 117, 2 (2015), 403–413.
- [82] Naomi Smith and Anne-Marie Snider. 2019. ASMR, affect and digitally-mediated intimacy. *Emotion, Space and Society* 30 (2019), 41–48.
- [83] Ho-Min Sohn. 2001. *The korean language*. Cambridge University Press.
- [84] Helen Spencer-Oatey and Peter Franklin. 2012. What is culture. *A compilation of quotations. GlobalPAD Core Concepts* 1 (2012), 22.
- [85] Brian H Spitzberg and William R Cupach. 2008. Fanning the flames of fandom: Celebrity worship, parasocial interaction, and stalking. *Stalking, threatening, and attacking public figures: A psychological and behavioral analysis* (2008), 287–321.
- [86] Charles Stangor, Rajiv Jhangiani, and Hammond Tarry. 2017. *Principles of social psychology*. BCcampus.
- [87] Jan-Philipp Stein, Priska Linda Breves, and Nora Anders. 2022. Parasocial interactions with real and virtual influencers: The role of perceived similarity and human-likeness. *New Media & Society* (2022), 14614448221102900.
- [88] Steven E Stern, John W Mullennix, and Ilya Yaroslavsky. 2006. Persuasion and social perception of human vs. synthetic voice across person as source and computer as source conditions. *International Journal of Human-Computer Studies* 64, 1 (2006), 43–52.
- [89] Meicheng Sun. 2020. K-pop fan labor and an alternative creative industry: A case study of GOT7 Chinese fans. *Global Media and China* 5, 4 (2020), 389–406.
- [90] Nurit Tal-Or and Yael Papirman. 2007. The fundamental attribution error in attributing fictional figures' characteristics to the actors. *Media psychology* 9, 2 (2007), 331–345.
- [91] Pulkit Tayal and Vijayakumar Bharathi. S. 2021. Reliability and trust perception of users on social media posts related to the ongoing COVID-19 pandemic. *Journal of Human Behavior in the Social Environment* 31, 1-4 (2021), 325–339.
- [92] Riva Tukachinsky and Gayle Stever. 2019. Theorizing development of parasocial engagement. *Communication Theory* 29, 3 (2019), 297–318.
- [93] Mary-Caitlyn Valentinsson. 2018. Stance and the construction of authentic celebrity persona. *Language in society* 47, 5 (2018), 715–740.
- [94] Will Verma, Pranshu & Oremus. 2023. AI voice clones mimic politicians and celebrities, reshaping reality. <https://www.washingtonpost.com/technology/2023/10/13/ai-voice-cloning-deepfakes/> Accessed: 2024-07-14.
- [95] Andrea J Vickery and Sabrina Ventrano. 2021. Listening goals and parasocial relationships: How listening styles impact the development of parasocial relationships with media personas. *International Journal of Listening* 35, 3

- (2021), 149–165.
- [96] Cynthia Vinney and Lisa A Vinney. 2017. That sounds familiar: The relationship between listeners’ recognition of celebrity voices, perceptions of vocal pleasantness, and engagement with media. *Journal of Radio & Audio Media* 24, 2 (2017), 320–338.
- [97] Hye-Sook Wang. 1995. The impact of family background on the acquisition of Korean honorifics. *Korean Language in America* (1995), 197–211.
- [98] Soyoung Wang and Seongcheol Kim. 2022. Users’ emotional and behavioral responses to deepfake videos of K-pop idols. *Computers in Human Behavior* 134 (2022), 107305.
- [99] Daniel J Weigel and Deborah S Ballard-Reisch. 2002. Investigating the behavioral indicators of relational commitment. *Journal of Social and Personal Relationships* 19, 3 (2002), 403–423.
- [100] Weverse. 2023. Weverse Shop - All Things for Fans! [https://weverseshop.io/en/shop/GL\\_USD/artists/2/sales/1506](https://weverseshop.io/en/shop/GL_USD/artists/2/sales/1506) Accessed: 2023-10-20.
- [101] Emily Wind. 2023. What’s behind the Taylor Swift friendship bracelets trend? <https://www.theguardian.com/music/2024/feb/07/taylor-swift-eras-tour-australia-friendship-bracelets-inspiration-beads-explained> Accessed: 2024-07-14.
- [102] Lilian Yang. 2022. K-pop apps create the illusion of private messaging with celebrities. <https://www.nngroup.com/articles/kpop-private-messaging/#:~:text=One%20way%20in%20which%20K,reply%20from%20many%20other%20fans> Accessed: 2023-09-10.
- [103] So-Yeon Yoon. 2021. [review] BTS makes learning Korean a dynamite experience. <https://korea JoongAngDaily Joins.com/2021/05/05/culture/books/Learn-Korean-with-TinyTAN-BTS/20210505154400339.html> Accessed: 2023-10-20.
- [104] Got7 Youngjae. 2021. Twitter Posting. [https://twitter.com/GOTYJ\\_Ars\\_Vita/status/1403598124073906183](https://twitter.com/GOTYJ_Ars_Vita/status/1403598124073906183) Accessed: 2021-08-30.

## A AI PRIVATE CALL INTERFACE OPTIONS

- (1) Date and time: Users could choose when they wanted to receive a simulated voice message.
- (2) The idol’s speech level: Users could choose how the message content of simulated voice message is structured by choosing one of three speech level options of chosen idols (formal, informal, any). Speech level refers to verb endings which are usually decided based on the relationship between speaker and hearer [97]. For example, when they choose ‘formal’ on the options, the script of the simulated voice message uses sentences which usually end with ‘ni-da’(니 다)’ or ‘yo’(요) that are used when speakers speak to a hearer who holds a higher position than them [97].
- (3) A call scenario: The options were ‘morning call’, ‘cheering up’, ‘birthday celebration’ and ‘daily chat.’ When they chose a scenario, they received a simulated voice message composed of one of several scripts for that scenario. The same scenarios were available for all idols.
- (4) The name to be called: Users could set the name that the voice message would call in the interface. If they set the name, they were able to listen to a voice message calling the name with synthetic voice of idols.

## B INTERVIEW QUESTIONS

- (1) Background
  - Which idols are you a fan of?
  - How would you describe your idol’s personality?  
(If participants are subscribing to several idols, ask for about all idols)
  - How do you usually communicate with idols?
  - How much do you usually know about AI
  - Can you explain what you know about how AI Private Call are made?
- (2) Impressions of using AI Private Call (Motivation / Expectation)
  - How long have you been using Universe?
  - Why do you subscribe to paid plan for Universe?
  - Why do you use AI Private Call?
  - How do you use AI Private Call?

- When do you mainly use AI Private Call?
  - Which scenario do you mainly use for an experience?
  - (if not currently being in use) What was your specific reason for not using AI Private Call?
  - (if not currently being in use) Are you unwilling to use it in the future?
- (3) Questions about voice message of AI Private Call
- Which idol do you usually choose when using AI Private Call?
  - How do you think about the message content of the voice message?
  - Do you feel that the message content of the AI private call was something the idol would actually say?
  - How is the voice in AI Private Call?
  - What do you think of when you hear AI Private Call?
  - How do you feel after using AI Private Call?
  - How did you feel towards your idols after using AI Private Call?
  - Have you ever shared your experience of using AI Private Call with others, if it is, could you share the experience?
  - Do you have any memorable AI Private Call conversations?
  - What was your favorite conversation?
  - What was the conversation you didn't like?
  - If AI Private Calls are improved, do you think you will use them?
  - Do you have any concerns about AI Private Call?
- (4) Questions about relationship with idols
- Do you feel like using AI Private Call creates a personal connection with your idol?
  - How will you feel about your idol after using AI Private Call?
  - Do you think you can feel the emotion of personal relationships with an idol by AI Private Call?
  - If using AI Private Call, will it affect the use of other services that idols communicate directly with?
- (5) Comparison with existing services
- (If participants have used one of social text message services for fan activities)  
What do you think is different from the social-text message service (e.g., Bubble, Private Message)?
  - (If participants have used live streaming platform for fan activities)  
What do you think is different from the live streaming platform?
- (6) Miscellaneous
- What is your opinion of AI trained with idol's data?
  - What are benefits of using AI Private Call?
  - What do you think is not good about private calls?
  - What improvements should be made to AI Private Call?

Received January 2024; revised July 2024; accepted October 2024