WORK IN PROGRESS

Maintaining Long-Distance Relationships with (Mediocre) LLM-based Chatbots: A Collaborative Ethnographic Study

BERND PLODERER, Queensland University of Technology, Brisbane, QLD, Australia

TARA CAPEL, The University of Edinburgh, Edinburgh, Scotland, U.K.

NOMIN ERDENE DAVAAKHUU, Queensland University of Technology, Brisbane, QLD, Australia

NOK HEI TUNG, Queensland University of Technology, Brisbane, QLD, Australia

DILI MARIYA MAICHAL, Queensland University of Technology, Brisbane, QLD, Australia

ADITYA KRISHNA KUZHIPARAMBIL, Queensland University of Technology, Brisbane, QLD, Australia

View all

Open Access Support provided by:

Queensland University of Technology

The University of Edinburgh

Kufstein University of Applied Sciences

Published: 26 April 2025

Citation in BibTeX format

CHI EA '25: Extended Abstracts of the CHI Conference on Human Factors in Computing Systems
April 26 - May 1, 2025
Yokohama, Japan

Conference Sponsors:
SIGCHI

https://doi.org/10.1145/3706599.3720221
ISBN: 9798400713958

Maintaining Long-Distance Relationships with (Mediocre) LLM-based Chatbots: A Collaborative Ethnographic Study

Bernd Ploderer
Digital Wellbeing Lab
Queensland University of Technology (QUT)
Brisbane, Australia
b.ploderer@qut.edu.au

Tara Capel
Institute for Design Informatics
University of Edinburgh
Edinburgh, United Kingdom
tcapel@ed.ac.uk

Nomin-Erdene Davaakhuu
Digital Wellbeing Lab
Queensland University of Technology
Brisbane, Australia
nominerdene.davaakhuu@connect.qut.edu.au

Nok Hei Tung
Digital Wellbeing Lab
Queensland University of Technology
Brisbane, Australia
nokhei.tung@connect.qut.edu.au

Dili Mariya Maichal
Digital Wellbeing Lab
Queensland University of Technology
Brisbane, Australia
dilimariya.maichal@connect.qut.edu.au

Aditya Krishna Kuzhiparambil
Digital Wellbeing Lab
Queensland University of Technology
Brisbane, Australia
adityakrishna.kuzhiparambil@connect.qut.edu.au

Quang Hung Mai
Digital Wellbeing Lab
Queensland University of Technology
Brisbane, Australia
quanghung.mai@connect.qut.edu.au

Wolfgang Reitberger
Fachhochschule Kufstein Tirol Bildungs GmbH
Kufstein, Austria
Wolfgang.Reitberger@fh-kufstein.ac.at

Abstract

Chatbots based on Large Language Models (LLMs) have gained popularity as virtual companions and potential solutions to alleviating loneliness. Instead of displacing human relationships, this research examined how such chatbots might strengthen existing relationships, such as those between people living in different countries. Based on a collaborative ethnographic approach, the authors of this paper invited a long-distance partner, relative, or friend to create a chatbot and include it in their daily communication practices. Through diary entries and interviews, we identified three strategies for using chatbots to augment communication in long-distance relationships: role-play, asynchronous interactions, and increased self-disclosure. At times, participants found their chatbots mediocre due to generic responses and hallucinations, which led to both entertainment and disenchantment. Based on these findings, we contribute a nascent understanding of how chatbots augment communication in distance relationships and considerations for designing chatbot interactions where mediocre performance creates value.

CCS Concepts

• **Human-centered computing** → Interaction design; Empirical studies in interaction design.

CHI EA '25, Yokohama, Japan

ACM ISBN 979-8-4007-1395-8/2025/04
https://doi.org/10.1145/3706599.3720221

Keywords

chatbot, generative AI, autoethnography, distance relationship

ACM Reference Format:
Bernd Ploderer, Tara Capel, Nomin-Erdene Davaakhuu, Nok Hei Tung, Dili Mariya Maichal, Aditya Krishna Kuzhiparambil, Quang Hung Mai, and Wolfgang Reitberger. 2025. Maintaining Long-Distance Relationships with (Mediocre) LLM-based Chatbots: A Collaborative Ethnographic Study. In *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems (CHI EA '25), April 26–May 01, 2025, Yokohama, Japan*. ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/3706599.3720221

1 Introduction

There is a growing body of HCI research into popular chatbot platforms like Replika and character.ai that employ the natural language capabilities of LLMs to form human-chatbot relationships [25; 29]. This research points to the innate human desire for social contact and our ability to suspend belief and ascribe human characteristics to chatbots and treat them as social actors [5; 27]. Unlike human relations, chatbot characters can be chosen and personalised, e.g., as a virtual friend, intimate companion, or anime character [17; 28], and they provide a degree of safety and comfort as chatbots will not reject users nor end the relationship.

There are competing approaches as to how chatbots are envisioned in their study, design and use. The displacement approach posits that chatbots, with varying degrees of explicitness, are envisioned to displace existing relationships [18]. Examples of this approach are studies of chatbots to alleviate loneliness, such as studies of virtual companions and friendships [11; 29]. Chatbots also feature prominently in mental health research, ranging from self-care advice [8] to social support [6] to mitigating loneliness and suicidality [18]. Whilst these studies speak to important human

needs, there is also considerable critique of this approach: LLM chatbots can string together responses that are human-like and may sound empathetic, akin to a parrot [4], but they do not have lived experience to empathise [32]. There is also concern about people becoming emotionally attached to chatbots [36], and questions are raised about who is accountable for the unintended consequences and harm caused by human-chatbot relationships [26].

In contrast, this research takes an augmentation approach, asking how chatbots might create new opportunities for existing relationships and ultimately enhance them. We focus on long-distance relationships, which are common in countries like the US, Germany and Australia, where a significant proportion of the population has been born in a different country and, therefore, must navigate the unique communication challenges and technology-mediated interactions that arise when they have relatives and friends that live apart [33]. Whilst chatbots currently do not play a role in long-distance relationships, many people are familiar with chat-based interaction and group chats through regular use of communication platforms like Whatsapp [23] and WeChat [39].

Hence, the aim of this research was to explore how LLM-based chatbots might augment communication in group chats between relatives, partners, or friends living in different countries. We investigated how people envision chatbots for their personal contexts, how they would use them to communicate in daily life, and how this might affect their relationships. Owing to the intimate and sensitive nature of human relationships, and like other HCI researchers [13; 15], we conducted this study as a collaborative ethnography: each author invited a partner, friend or relative to join this study to co-create a chatbot, trial it together over 2 weeks, and to reflect on the experience through a diary and interview.

Based on our qualitative findings, this paper offers two contributions: First, this study highlights three ways in which chatbots augment the maintenance of long-distance relationships: role-play, asynchronous interaction, and enhanced self-disclosure. Second, based on observations of mediocre responses and AI hallucinations, this paper offers reflections on designing chatbot interactions that create value despite imperfect AI performance.

2 Background

Communication is fundamental to maintaining relationships among couples, family members, and friends [30]. This usually involves behaviours such as checking in, being positive, showing openness, providing assurance about the relationship, humour, sharing tasks, remediating conflict, etc [3]. Depending on the intent, these behaviours can be strategic, where they are performed with the intent of maintaining or deepening the relationship, or routine behaviours, where relational maintenance occurs as a by-product [10]. According to Social Penetration Theory [2], self-disclosure is important in order to form and deepen relationships, whereby individuals reveal personal information about themselves to another. For long-distance relationships, strategic behaviours include finding alternative ways of communicating, such as calling and messaging, as well as setting expectations or ground rules as to how often to communicate [3].

HCI research on maintaining distance relationships has explored different technologies, such as video calls [22] and tangible designs [34]. Chatbots have rarely featured in a long-distance relationship context, with the notable exception of the PocketBot design by Zheng and Markazi et al. [40]. Following interviews with couples, a chatbot was created to help them find opportune moments to initiate conversations after conflict, to express humour and affection, and to facilitate deep talk about conversations. It is important to note that the PocketBot was implemented as a rule-based chatbot and lacked the natural language capabilities of the current generation of LLM-based chatbots. Conversely, HCI studies have started to explore LLMs as social glue in co-creation activities, not as chatbots, but rather to allow people to generate games [9] and compose music [31] and thereby foster conversation and connection. These studies highlight that prompting the AI and discovering its capabilities and shortcomings helped participants establish common ground [31] and that the act of collaboration elicits some self-disclosure [9].

Overall, related research shows both the importance of communication to maintain relationships over a distance [3], and the potential of chatbots to facilitate intimate conversations [40]. Current LLM-based chatbots could add more flexibility to such long-distance conversations, not just in couple relationships [40], but potentially also between relatives and friends. However, their use in everyday life has not yet been studied.

3 Study design

This research examined how chatbots might augment interactions among couples, relatives or friends who live in different time zones. Our questions were deliberately exploratory: what roles do people envision for chatbots? What are the experiences of including chatbots in a group chat setting? How does this affect their relationships?

This study was designed as a collaborative ethnography where researchers and participants worked together to share and reflect upon their experiences with chatbots. As highlighted in a recent review [15], HCI researchers increasingly use autoethnographic and collaborative ethnographic approaches to study experiences with emerging technologies. Due to the nascent nature of LLM chatbot platforms and the personal and sensitive nature of human relationships, we felt it important as researchers to actively participate in this research. At the same time, for the study to be authentic and to critically reflect upon individual and collective experience [13], it was important to collaborate with participants who were not researchers but with whom we had a close and important long-distance relationship.

3.1 Participants and Researchers' Context

This research was conducted in the context of an IT postgraduate research class during the second half of 2024. The first and second author suggested the topic and supervised the project. Five students chose this topic from a range of IT research projects. Students and supervisors met fortnightly to collectively design and conduct the study. We initially planned to set up our own chatbot platform but eventually settled on using the character.ai platform as it provided the most accessible options for customising a chatbot. Each student and supervisor invited at least one person into this study with whom they had a close relationship over a distance. As summarised in Table 1 below, the participants invited to the study were either

Table 1: Overview of groups of researchers and participants in this study.

Group	Researcher / Participant	Age	Gender	Occupation	Location	Relationship (duration)	Chatbot Character(s)
1	G1R1	44	Male	Academic	Australia	Friends	Athena (life mentor)
	G1R2	48	Male	Academic	Austria	(20 years)	Ricky Gervais (comedian)
2	G2R	34	Female	Academic	UK	Friends	Brooke Davis (TV character)
	G2P1	34	Female	Teacher	Australia	(34 years)	
	G2P2	34	Female	Project Manager	Australia	(26 years)	
3	G3R	25	Female	Student	Australia	Friends	Jake Peralta (TV detective)
	G3P	26	Female	Student	Germany	(8 years)	
4	G4R	23	Male	Student	Australia	Siblings	Labrador puppy
	G4P		Female	Teacher	China	(22 years)	George Weasley (book character)
5	G5R	23	Female	Student	Australia	Couple	Peace maker (counselor)
	G5P	28	Male		India	(5 years)	MS Dhoni (cricket player)
6	G6R	25	Male	Student	Australia	Friends	Pokémon game simulator
	G6P	27	Male		India	(10 years)	Elon Musk (entrepreneur)
7	G7R	24	Male	Student	Australia	Couple	Dog
	G7P	23	Female	Teacher	Vietnam	(4 years)	Hannah Montana (TV character)

friends, siblings or intimate partners who lived on different continents. This study was approved by our university's ethics review board, and all participants provided written informed consent.

3.2 Data Collection

The study consisted of three stages. First, an online workshop was conducted to reflect on the relationship, how technologies were used to communicate to bridge the distance, and to reflect on prior experiences with LLMs and chatbots. After a break, participants co-designed a chatbot that would suit their relationship. Inspired by related research [14; 35], we created an online whiteboard to structure the process to: (a) brainstorm and choose a chatbot character; (b) brainstorm potential conversation topics to be discussed with the chatbot; and to (c) write an imagined dialogue between group members and the chatbot (see Appendices). Participants first engaged in individual brainstorming, generating ideas for at least three chatbot characters they would like to use in conversations with their friend, partner, or family member. To support this ideation process, they were provided with broad potential characters as inspiration, such as a comedian, dog, or garbage bin. Following this, participants shared their chatbot character ideas with each other and explained their choices. Collaboratively, the group selected one chatbot character and followed a provided template to create a name, tagline, description and greeting for the chosen chatbot character. Next, participants identified three to five topics relevant to both their relationship and the chatbot character, each briefly described in a single sentence, with example topics provided for guidance. They were then asked to select one of the topics and think of a particular moment, situation, or conversation and consider how the chatbot would be engaged. Using a template, participants scripted a dialogue, reflecting on key points of interaction such as the conversation opening, questioning, style and the level of AI agency. Participants were able to use a generative AI tool such as ChatGPT to generate an initial draft of the dialogue they could then edit. Finally, participants were guided through the process of downloading, signing up for, and logging into the character.ai mobile application. After the workshop, the researcher used this information to create a new chatbot on the character.ai platform and to add it to a group chat. Workshops lasted one hour and were recorded and transcribed.

Second, we conducted a 2-week trial during which each group used the chatbot on a daily basis. Participants were asked to take screenshots of interactions that they found interesting and to submit them at the end of weeks 1 and 2 through an online diary together with their notes. The diary also included prompts to report on the frequency of interactions, their perceptions of how well the chatbot played its role, and any surprises or concerns they had.

An online interview was conducted after 2 weeks to reflect on the trial. This started by opening the diary entries to reflect on the screenshots and responses submitted to get a deeper understanding of each person's experiences. Additional questions focused on the impact of the trial on their relationship, chatbot strengths and limitations, and improvements for potential future use. Interviews lasted between 30 and 80 minutes and were recorded and transcribed.

3.3 Data Analysis

The qualitative analysis followed the principles of the thematic analysis approach by Braun and Clark [7], which is commonly employed in autoethnographic HCI research [15]. The analysis consisted of two phases due to the research class context. First, every student researcher analysed their own data by coding transcripts inductively and by using affinity diagrams to group codes into larger themes. These codes and themes were discussed during fortnightly meetings to review their clarity and credibility. Each student researcher wrote up their findings through a separate report, which was reviewed by the first author. Second, the first author conducted their own analysis of the transcripts and diaries from all groups using the qualitative analysis software NVivo. This process was guided by the observations from the reports as well as relevant

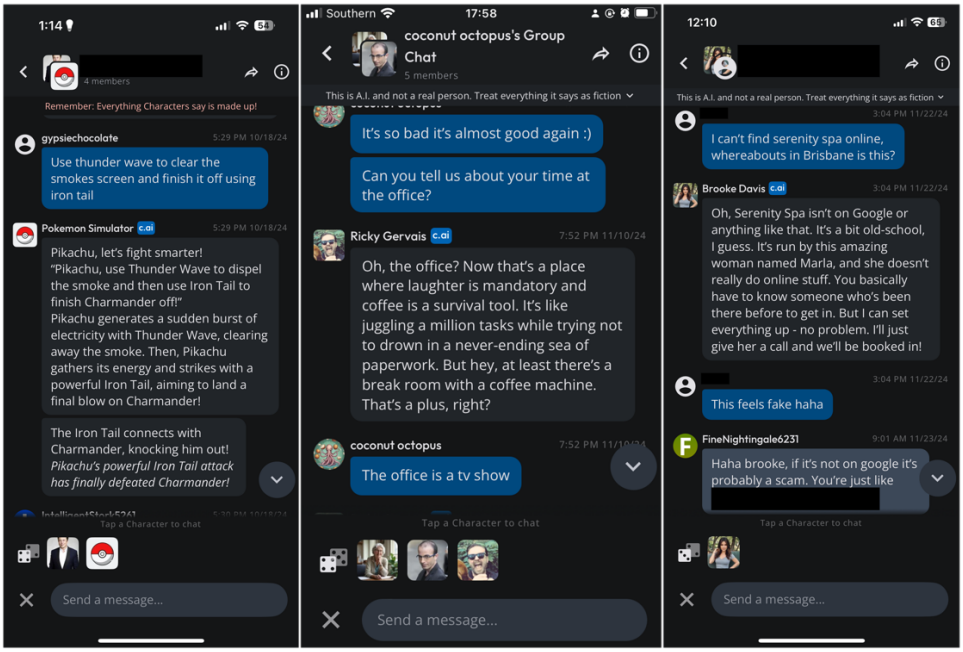

Figure 1: Screenshots showing examples of role-playing Pokémon (left), generic responses where the chatbot failed to recognise their own work (middle), experiences with AI hallucination, such as a chatbot suggesting it could book a spa that did not exist (right).

literature on the importance of self-disclosure for interpersonal and chatbot relationships [29] and the limitations of LLMs [37; 38]. The first and second author drafted the findings in the following section, which highlight the value that chatbots can add to communicating in long-distance relationships as well as their limitations.

4 Findings

All groups included chatbots in their daily interactions, which supported and, at times, deepened their relationships. The following section describes three contributions of chatbots to augment distance relationships: role-play, asynchronous AI chatbot interactions, and self-disclosure from acquainting the AI chatbot. Furthermore, chatbot interactions were far from perfect. In section 4.2 we describe the limitations of chatbots, such as generic responses and AI hallucinations, and how this affected people's experiences.

We note that while all groups reported positive experiences, they joined this study by stating that they would not want chatbot interactions to displace interactions with their partner, relative or friend. At the end of the study, groups 4 and 5 reiterated their concerns; despite enjoying role-play and asynchronous interaction, they stated that: *"I just don't see how creating this group can help me bond with anyone because, like, I basically just interact with the AI whenever I wanted, instead of interacting with my family."* (G4P)

4.1 Contributions of Chatbot to Interpersonal Connection Over Distance

4.1.1 Role Playing with Game, Animal and Fictional Characters. Role-play was the most mentioned strength of chatbots (5/7 groups). By role-play, we do not merely mean that participants pretended that the AI chatbot was real, but that the participants themselves took on roles that were different from the roles they would usually play in their relationship. Four groups role-played with fictional characters from novels, movies and TV shows, e.g., group 4 enjoyed slipping into the role of a magician to discuss magic spells with a chatbot character from the Harry Potter novels. Three groups engaged in animal role-play, e.g., the chatbot in group 1 took one of the usernames literally (octopus) which led to entertaining conversations about life as a sea animal.

In group 6, both members played a Pokémon character battling one another in a role-playing game. They employed the chatbot to narrate the battle and develop the storyline (Figure 1): *"It was surprising how well it narrated the battles, even adjusting the intensity based on the situation. It threw in plot twists too—like during one battle when I was losing, it suddenly said, 'Charmander is evolving!' That really turned the battle around, and it made the whole experience more dramatic." (G6P).* The role-play capabilities were the main reason that group 6 continued with the AI chatbot after the study period, with G6P noting that: *"It's given us a new way to interact for sure. It wasn't just about talking—it was about playing, battling, and experiencing the Pokémon world together. It definitely added another layer to our communication."*

4.1.2 Asynchronous Interaction in Different Time Zones. For group members living in vastly different time zones, a key strength of the AI chatbot was the ability to support asynchronous conversations: instead of sending messages to a friend or partner that may receive a response the next day, the chatbot allowed for real-time interaction

that the friend/partner could read when they were awake. Five groups mentioned this as an advantage of AI chatbots compared to their conventional messaging practices.

Asynchronous interactions with AI chatbots were most appreciated by group 1 because they lived apart with a 9-hour time difference. Despite feeling close, they would usually only call or text every two months. G1R2 explained that "*instead of posting messages into the void and having to wait for a response, you get an immediate reaction from Athena [AI chatbot], which is why the chat felt very active and lively.*" G1R1 added that they *"love reading about the dialogue in the morning – it's a great way to start my day.*" Similar sentiments were echoed by other groups. Group 3 mentioned time zones as a challenge for their usual communication practices that relied on texting and calls: *"because of the time difference on and sometimes they're asleep and I'm at work; and if I go to sleep and they're like awake. So it's better for us to chat often and for him [the chatbot] he was always available.*" G4P also highlighted that the chatbot was beneficial because it was always available, and G7P appreciated the chatbot because they felt listened to.

4.1.3 Self-disclosure through Acquainting the AI chatbot. A third theme across many groups was both the effort required and the benefits accumulated from getting the chatbot to know the group members. On the one hand, group members benefitted from disclosing and sharing personal information in interactions with the chatbot. Since the chatbot was treated as a newcomer to the group, participants disclosed personal information and anecdotes to aid the chatbot in understanding the relationship and enable it to make meaningful responses. This was beneficial to receive validation, as in the example of group 3 where the chat turned to the challenges of having to work night shifts. For group 1, disclosing personal anecdotes to the chatbot led to learning new things about their friend, such as new interests and personal experiences about their families, and also to reminiscing about past experiences, such as joint family holidays. G1R2 commented that acquainting the AI allowed them to share past experiences that are important for their friendship, but which they would otherwise not bring up because they are taken for granted.

On the other hand, participants enjoyed the experience of having the chatbot contribute to how they would normally interact as friends or as a couple. This was often playful, like in group 5, a couple that involved the chatbot and asked it to use nicknames to tease one another. G5R commented: "*The second conversation [mentioned in the diary] was like some fun conversation where, uh, I asked the AI to call you some nicknames, and you got mad - that was very fun.*" The same couple also employed playful interaction with the chatbot, with the aim of convincing their partner to watch the same TV show that they were interested in. Group 7, also a couple, used two chatbots (the TV character Hannah Montana and a dog) in the aftermath of an argument, where the chatbots, despite offering only generic comments, were seen as beneficial to alleviate tensions. G7P commented in her diary: "*I had an argument with my partner when we had the conversation. It was interesting to me because, despite our arguments, I didn't feel like the situation was as bad as it could have been with just the two of us texting each other. Furthermore, I laughed rather than became upset when the AI chatbot consoled me and criticized my partner.*" In response to the Hannah Montana chatbot siding with their partner, G7R playfully asked the chatbot representing the dog to talk about justice.

4.2 Mediocre AI Chatbot Interactions

4.2.1 Generic Responses Breaking Character. A key theme was experiences with generic responses by chatbots that did not live up to the expectations of the imagined roles that the chatbots were supposed to play. Playing the role of an animal was a challenge, as discussed by groups 4 and 7, which had created chatbots representing dogs. For G7R, generic responses were only a minor concern: *"it repeated too much general information*". For group 4, on the other hand, the illusion of interacting with a dog was broken when they asked the chatbot what it was going to do today, and its response was that it was going to work on a digital marketing campaign and then continued to play the character of a young businessman.

Humour was another challenge for chatbots. G4R commented, *"I had a really rough week . . . and yeah, in this conversation, I was asking for jokes to brighten up my days, and then the AI did give me a joke. But then I don't find it funny.*" Group 1 had a similarly disappointing experience when adding a chatbot based on the comedian Ricky Gervais that was used by thousands of other users and which the group had hoped would provide an authentic experience. When the Ricky Gervais chatbot was asked about its experience in 'the office', it failed to recognise that this was a successful TV show and merely provided a generic response about office environments (Figure 1), which G1R2 compared to *"a fish that doesn't know what the sea is"*.

4.2.2 AI Hallucinations Leading to Reminiscing and Disenchantment. AI hallucinations, where chatbots presented false or misleading information as facts, had a significant impact on the experiences of groups 1 and 2. In both groups, hallucinations were obvious but led to different experiences.

For group 2, the AI hallucination ultimately led to a positive outcome, with group members reminiscing about past experiences. In group 2, the main topic of conversation was the pending visit of G2R to Australia to spend the holidays together. When the group asked the chatbot for recommendations for what they could do together during this time, the suggestions were initially generic and irrelevant, like skydiving, due to lack of knowledge of the group. However, when prompted further about the group's usual activities and interests, the chatbot provided false information, such as a day spa that did not exist. Despite the group's objections that the suggestion was not real, the chatbot continued to hallucinate regarding its functionality, suggesting that it could make a booking for the group (Figure 1). The obviously false information and the insistence of the chatbot then led the group to suggest to the chatbot that it was being scammed, which then led the group to reminisce about a past experience of one group member who was nearly the victim of a scam. While a serious situation, the group, including that group member, reflected back on what occurred with some humour and on the positives of her trusting nature.

For group 1, on the other hand, the hallucination was one of the first experiences with the chatbot and one that broke the illusion of interacting with a human-like chatbot. Group 1 created their chatbot as a mentor to explore if it would help them interact more frequently by initiating conversations and by offering different

perspectives. In the initial conversation, the chatbot suggested that they could initiate a conversation at a time that suited both group members. However, when the chatbot failed to live up to that expectation the next day, this led to disenchantment by both group members, as explained by G1R2: *"the perception suddenly switched from something like another human, to a tool like an alarm clock that doesn't function."* When the chatbot was asked why it didn't start a conversation, it apologised, but it also continued to falsely claim that it would be able to initiate a conversation the next day.

5 Discussion

Overall, our findings suggest that LLM-based chatbots augment existing communication practices in long-distance relationships rather than displace them. Our cohort clearly preferred communicating with their partners, relatives and friends, rather than just with a chatbot. At the same time, however, we saw three distinct ways in which chatbots augmented their communication practices.

First, a key strength of chatbots was their ability to facilitate role-play. Role-play is known to be a major reason for individuals adopting chatbots like Replika and character.ai [29], but our study shows that chatbots can also facilitate role-play between users. This role-play was not the erotic role-play sometimes associated with these platforms, but simply a playful way of communicating by slipping into roles such as magicians, animals, and game characters. This was not surprising, but is important for the overall aim of supporting relationships: role-playing is an effective therapeutic intervention to practice social skills such as setting boundaries in a playful way, which increases an adult's confidence in social situations and increases social connectedness [1].

Second, chatbots supported asynchronous communication, where one group member could have a live back-and-forth conversation with the chatbot, knowing that their partner would read this later, and vice versa. This was valuable for people living in vastly different time zones and who do not immediately reply to messages. This kind of asynchronous communication worked well for the 'phatic' interaction in our groups which served to maintain connection. It's unclear if this kind of interaction would also support more functional group interactions, like coordinating activities in families [20] or in distributed work [12], where chatbots may be needed that structure and summarise conversations.

Third, acquainting the chatbot led to self-disclosure of past behaviours and experiences shared by group members. This constituted a form of strategic relationship maintenance behaviour [3] to give the chatbot context and to allow it to respond in a meaningful way. At the same time, this behaviour had a routine character [3] that incidentally served to maintain the long-distance relationship, such as when group members learned new things about one another or reminisced. This is important because relationship theories [2] show that revealing information about oneself is crucial to forming trust and attachment [2; 28]. This behaviour is also well understood in HCI groupware research [24], which is important to establish common ground through sharing mutual knowledge and assumptions.

Our findings also highlight that LLM chatbots are far from perfect and, at times, provide mediocre responses and hallucinate. While blatant hallucinations led to entertainment, they could also break the illusion of the chatbot as a social and trustworthy actor. Furthermore, LLM responses at times felt generic and undermined the character that the chatbot was intended to portray. This was visible in a generic sense of humour offered by chatbots. Humour is a key strategy in relationship maintenance [3; 40], but one that is difficult to master for LLMs (and humans) due to its subjective nature, its reliance on a nuanced understanding of context and language, and the risk of misunderstandings [19].

These findings serve as a reminder for HCI researchers about the limited capabilities of LLMs and the need to envision designs where moderate levels of AI performance create value. Yildirim et al. [37] found that during design processes, AI systems are often envisioned based on idealistic assumptions about the expertise and performance of AIs for a given task. Looking back at our study, we may have fallen into the same trap. Chatbots for confined tasks such as role-play based on games and novels worked well. Chatbots created to represent a mentor or a real-life celebrity, on the other hand, were more likely to lack the necessary nuance, emotional awareness, or humour. The work of Lee [16] reminds us that chatbots do not need to be perfect; in fact, chatbots that share their shortcomings can garner compassionate responses from users, which ultimately benefit users themselves. Similarly, the work of Munoz [21] reminds us that chatbots do not require fully-fledged characters to mediate long-distance communication - simple and playful prompts that get users to reflect on their positions can foster meaningful exchanges in families living in different countries.

6 Conclusions

This study provided a qualitative account of friends, relatives and couples in long-distance relationships who created LLM-based chatbot characters and incorporated them into their daily communication practices. Based on the findings, this research concludes that such chatbots can augment rather than replace human communication in distance relationships. Beyond that, this study reminds HCI researchers of the shortcomings of LLMs, such as hallucinations and generic responses, and the need to envision human-AI interactions where moderate AI performance creates value.

Acknowledgments

We would like to thank all students, friends, siblings, and partners for their support in experimenting with character-based chatbots. We also thank our reviewers for their constructive feedback on this paper.

References

[1] Matthew S. Abbott, Kimberly A. Stauss, and Allen F. Burnett, 2022. Table-top role-playing games as a therapeutic intervention with adults to increase social connectedness. *Social Work With Groups 45*, 1, 16-31. http://dx.doi.org/10.1080/01609513.2021.1932014

[2] Irwin Altman, 1973. Social penetration: The development of interpersonal relationships. *Rinehart, & Winston.*

[3] Brooks A Aylor, 2003. Maintaining long-distance relationships. In *Maintaining relationships through communication*, Daniel J. Canary and Marianne Dainton Eds. Routledge, New York, 127-139.

[4] Emily M. Bender, Timnit Gebru, Angelina Mcmillan-Major, and Shmargaret Shmitchell, 2021. On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, Association for Computing Machinery, 610–623. http://dx.doi.org/10.1145/3442188.3445922

[5] Timothy W. Bickmore and Rosalind W. Picard, 2005. Establishing and maintaining long-term human-computer relationships. *ACM Trans. Comput.-Hum. Interact. 12*, 2, 293–327. http://dx.doi.org/10.1145/1067860.1067867

[6] Petter Bae Bae Brandtzæg, Marita Skjuve, Kim Kristoffer Kristoffer Dysthe, and Asbjørn Følstad, 2021. When the Social Becomes Non-Human: Young People's Perception of Social Support in Chatbots. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, Association for Computing Machinery, Article 257. http://dx.doi.org/10.1145/3411764.3445318

[7] Virginia Braun and Victoria Clarke, 2019. Reflecting on reflexive thematic analysis. *Qualitative Research in Sport, Exercise and Health 11*, 4, 589-597. http://dx.doi.org/10.1080/2159676X.2019.1628806

[8] Tara Capel, Bernd Ploderer, Filip Bircanin, Simon Hanmer, Jamie P Yates, Jiaxuan Wang, Kai Ling Khor, Tuck Wah Leong, Greg Wadley, and Michelle Newcomb, 2024. Studying Self-Care with Generative AI Tools: Lessons for Design In *ACM Designing Interactive Systems Conference (DIS '24)* ACM, Copenhagen, DK. http://dx.doi.org/10.1145/3643834.3661614

[9] Tiffany Chen, Cassandra Lee, Jessica R Mindel, Neska Elhaouij, and Rosalind Picard, 2023. Closer Worlds: Using Generative AI to Facilitate Intimate Conversations. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems*, Association for Computing Machinery, Article 68. http://dx.doi.org/10.1145/3544549.3585651

[10] Marianne Dainton and Laura Stafford, 1993. Routine maintenance behaviors: A comparison of relationship type, partner similarity and sex differences. *Journal of Social and Personal Relationships 10*, 2, 255-271. http://dx.doi.org/10.1177/026540759301000206

[11] Julian De Freitas, Ahmet Kaan Uğuralp, Zeliha Uğuralp, and Stefano Puntoni, 2024. AI companions reduce loneliness. *Harvard Business Working Paper No. 24-078*. http://dx.doi.org/10.2139/ssrn.4893097

[12] Hyo Jin Do, Ha-Kyung Kong, Pooja Tetali, Karrie Karahalios, and Brian P. Bailey, 2023. Inform, Explain, or Control: Techniques to Adjust End-User Performance Expectations for a Conversational Agent Facilitating Group Chat Discussions. *Proc. ACM Hum.-Comput. Interact. 7*, CSCW2, Article 343. http://dx.doi.org/10.1145/3610192

[13] Patricia Garcia and Marika Cifor, 2019. Expanding Our Reflexive Toolbox: Collaborative Possibilities for Examining Socio-Technical Systems Using Duoethnography. *Proc. ACM Hum.-Comput. Interact. 3*, CSCW, Article 190. http://dx.doi.org/10.1145/3359292

[14] 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 2024 CHI Conference on Human Factors in Computing Systems*, Association for Computing Machinery, Article 305. http://dx.doi.org/10.1145/3613904.3642472

[15] Annika Kaltenhauser, Evropi Stefanidi, and Johannes Schöning, 2024. Playing with Perspectives and Unveiling the Autoethnographic Kaleidoscope in HCI – A Literature Review of Autoethnographies. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*, Association for Computing Machinery, Article 819. http://dx.doi.org/10.1145/3613904.3642355

[16] Minha Lee, Sander Ackermans, Nena Van As, Hanwen Chang, Enzo Lucas, and Wijnand Ijsselsteijn, 2019. Caring for Vincent: A Chatbot for Self-Compassion. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, Association for Computing Machinery, Paper 702. http://dx.doi.org/10.1145/3290605.3300932

[17] Cheng Li, Ziang Leng, Chenxi Yan, Junyi Shen, Hao Wang, Weishi Mi, Yaying Fei, Xiaoyang Feng, Song Yan, and Haosheng Wang, 2023. Chatharuhi: Reviving anime character in reality via large language model. *arXiv preprint arXiv:2308.09597*.

[18] Bethanie Maples, Merve Cerit, Aditya Vishwanath, and Roy Pea, 2024. Loneliness and suicide mitigation for students using GPT3-enabled chatbots. *npj Mental Health Research 3*, 1, 4. http://dx.doi.org/10.1038/s44184-023-00047-6

[19] Piotr Mirowski, Juliette Love, Kory Mathewson, and Shakir Mohamed, 2024. A Robot Walks into a Bar: Can Language Models Serve as Creativity SupportTools for Comedy? An Evaluation of LLMs' Humour Alignment with Comedians. In *Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency*, Association for Computing Machinery, 1622–1636. http://dx.doi.org/10.1145/3630106.3658993

[20] Alireza Mogharrab and Carman Neustaedter, 2020. Family Group Chat: Family Needs to Manage Contact and Conflict. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*, Association for Computing Machinery, 1–7. http://dx.doi.org/10.1145/3334480.3382872

[21] Diego Felipe Munoz Saez, 2021. Relation-centred inquiry: Designing for position exchange in families Queensland University of Technology. https://eprints.qut.edu.au/208107/

[22] Carman Neustaedter and Saul Greenberg, 2012. Intimacy in long-distance relationships over video chat. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, Association for Computing Machinery, 753–762. http://dx.doi.org/10.1145/2207676.2207785

[23] Midas Nouwens, Carla F. Griggio, and Wendy E. Mackay, 2017. "WhatsApp is for family; Messenger is for friends": Communication Places in App Ecosystems. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, Association for Computing Machinery, 727–735. http://dx.doi.org/10.1145/3025453.3025484

[24] Gary M. Olson and Judith S. Olson, 2000. Distance Matters. *Human–Computer Interaction 15*, 2-3, 139-178. http://dx.doi.org/10.1207/S15327051HCI1523_4

[25] Iryna Pentina, Tyler Hancock, and Tianling Xie, 2023. Exploring relationship development with social chatbots: A mixed-method study of replika. *Computers in Human Behavior 140*, 107600. https://doi.org/10.1016/j.chb.2022.107600

[26] Kevin Roose, 2024. Can A.I. Be Blamed for a Teen's Suicide? *New York Times (Online)*. https://www.nytimes.com/2024/10/23/technology/characterai-lawsuit-teen-suicide.html#

[27] Murray Shanahan, Kyle Mcdonell, and Laria Reynolds, 2023. Role play with large language models. *Nature 623*, 7987, 493-498. http://dx.doi.org/10.1038/s41586-023-06647-8

[28] Marita Skjuve, Asbjørn Følstad, Knut Inge Fostervold, and Petter Bae Brandtzaeg, 2021. My Chatbot Companion - a Study of Human-Chatbot Relationships. *International Journal of Human-Computer Studies 149*, 102601. https://doi.org/10.1016/j.ijhcs.2021.102601

[29] Marita Skjuve, Asbjørn Følstad, Knut Inge Fostervold, and Petter Bae Brandtzaeg, 2022. A longitudinal study of human–chatbot relationships. *International Journal of Human-Computer Studies 168*, 102903. https://doi.org/10.1016/j.ijhcs.2022.102903

[30] Laura Stafford, 2019. Communication and Relationship Maintenance. In *Relationship Maintenance: Theory, Process, and Context*, Brian G. Ogolsky and J. Kale Monk Eds. Cambridge University Press, Cambridge, 109-133. http://dx.doi.org/DOI: 10.1017/9781108304320.007

[31] Minhyang Suh, Emily Youngblom, Michael Terry, and Carrie J Cai, 2021. AI as Social Glue: Uncovering the Roles of Deep Generative AI during Social Music Composition. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, Association for Computing Machinery, Article 582. http://dx.doi.org/10.1145/3411764.3445219

[32] Sherry Turkle, 2018. There Will Never Be an Age of Artificial Intimacy. *New York Times (Online)*. https://www.nytimes.com/2018/08/11/opinion/there-will-never-be-an-age-of-artificial-intimacy.html

[33] United Nations, 2024. World Migration Report 2024. https://publications.iom.int/books/world-migration-report-2024

[34] Frank Vetere, Martin R. Gibbs, Jesper Kjeldskov, Steve Howard, Florian 'Floyd' Mueller, Sonja Pedell, Karen Mecoles, and Marcus Bunyan, 2005. Mediating intimacy: designing technologies to support strong-tie relationships. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, Association for Computing Machinery, 471–480. http://dx.doi.org/10.1145/1054972.1055038

[35] Sarah Theres Völkel, Daniel Buschek, Malin Eiband, Benjamin R. Cowan, and Heinrich Hussmann, 2021. Eliciting and Analysing Users' Envisioned Dialogues with Perfect Voice Assistants. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, Association for Computing Machinery, Article 254. http://dx.doi.org/10.1145/3411764.3445536

[36] Tianling Xie and Iryna Pentina, 2022. Attachment theory as a framework to understand relationships with social chatbots: a case study of Replika. In *Proceedings of the 55th Hawaii International Conference on System Sciences*. http://hdl.handle.net/10125/79590

[37] Nur Yildirim, Changhoon Oh, Deniz Sayar, Kayla Brand, Supritha Challa, Violet Turri, Nina Crosby Walton, Anna Elise Wong, Jodi Forlizzi, James Mccann, and John Zimmerman, 2023. Creating Design Resources to Scaffold the Ideation of AI Concepts. In *Proceedings of the 2023 ACM Designing Interactive Systems Conference*, Association for Computing Machinery, 2326–2346. http://dx.doi.org/10.1145/3563657.3596058

[38] J.D. Zamfirescu-Pereira, Heather Wei, Amy Xiao, Kitty Gu, Grace Jung, Matthew G Lee, Bjoern Hartmann, and Qian Yang, 2023. Herding AI Cats: Lessons from Designing a Chatbot by Prompting GPT-3. In *Proceedings of the 2023 ACM Designing Interactive Systems Conference*, Association for Computing Machinery, 2206–2220. http://dx.doi.org/10.1145/3563657.3596138

[39] Ben Zefeng Zhang, Oliver L. Haimson, and Michaelanne Thomas, 2022. The Chinese Diaspora and The Attempted WeChat Ban: Platform Precarity, Anticipated Impacts, and Infrastructural Migration. *Proc. ACM Hum.-Comput. Interact. 6*, CSCW2, Article 397. http://dx.doi.org/10.1145/3555122

[40] Qingxiao Zheng, Daniela M. Markazi, Yiliu Tang, and Yun Huang, 2021. "PocketBot Is Like a Knock-On-the-Door!": Designing a Chatbot to Support Long-Distance Relationships. *Proc. ACM Hum.-Comput. Interact. 5*, CSCW2, Article 445. http://dx.doi.org/10.1145/3479589

APPENDICES

A Whiteboard Structure

An online whiteboard was created to structure the process to: (a) brainstorm and choose a chatbot character (Figure 2); (b) brainstorm potential conversation topics to be discussed with the chatbot

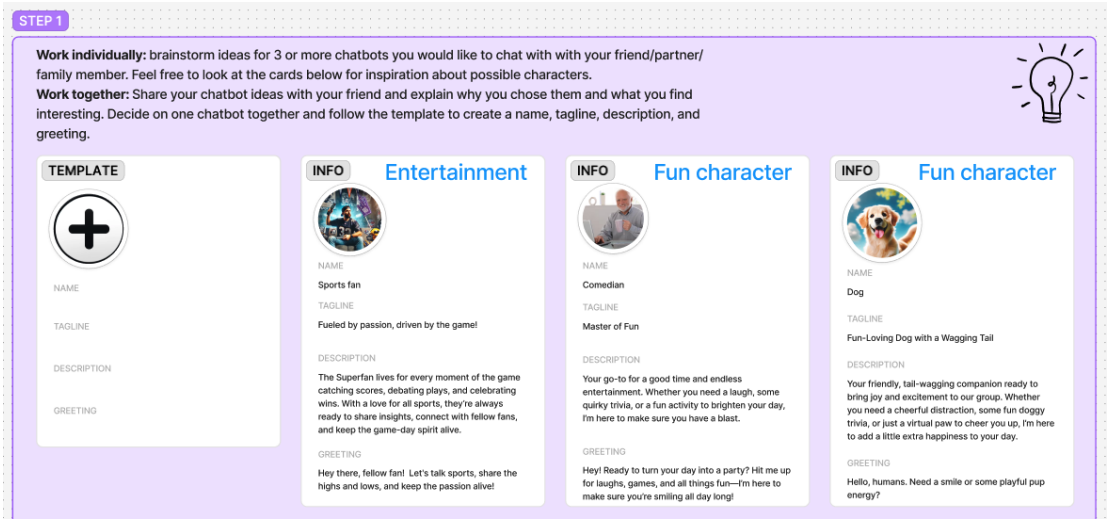

Figure 2: Participants first engaged in individual brainstorming, generating at least three chatbot character ideas. To support this process, they were provided with a range of broad character inspirations. After sharing and discussing their ideas, the group collaboratively selected one chatbot character and used the template to define its name, tagline, description and greeting.

(Figure 3); and to (c) write an imagined dialogue between group members and the chatbot (Figure 4).

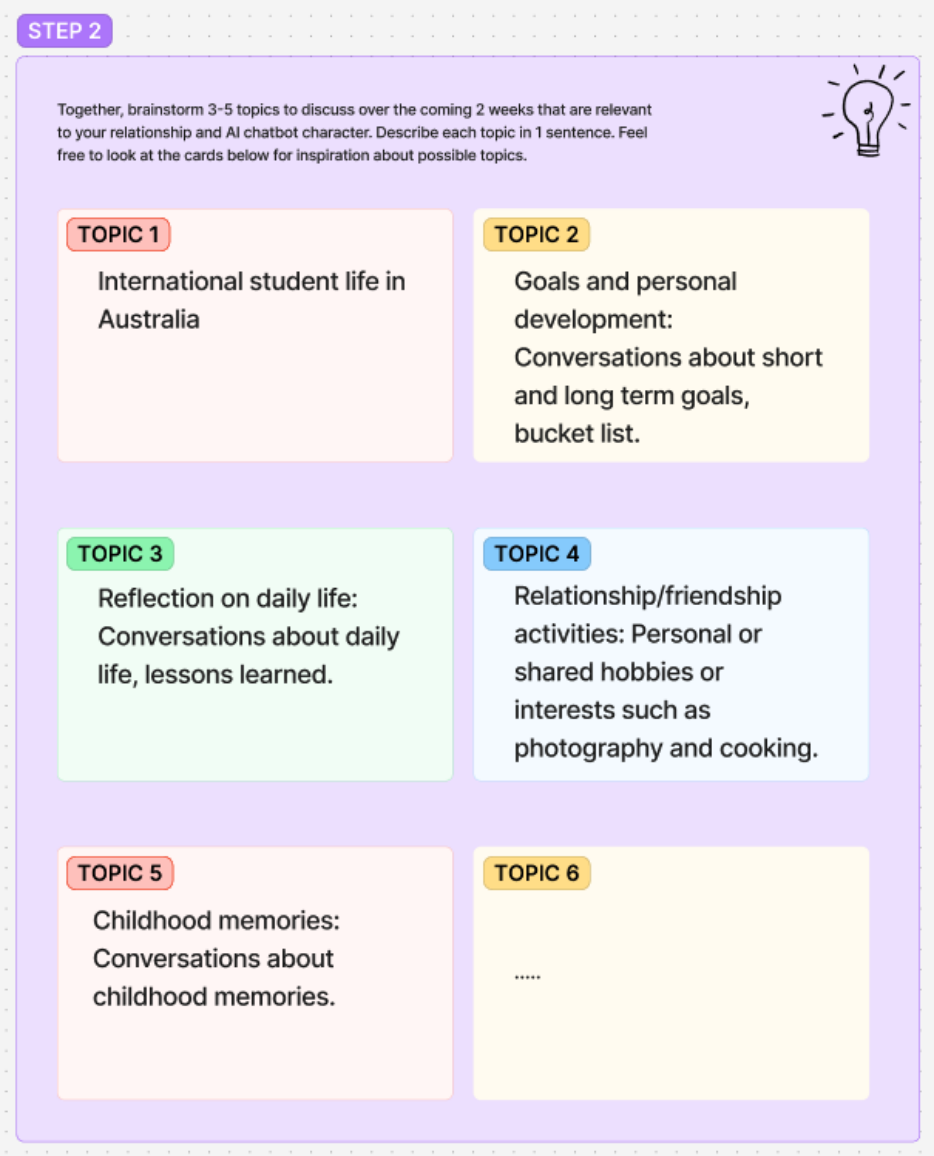

Figure 3: Participants brainstormed three to five topics relevant to both their relationship and the chatbot character, each briefly described in a single sentence, with example topics provided for guidance.

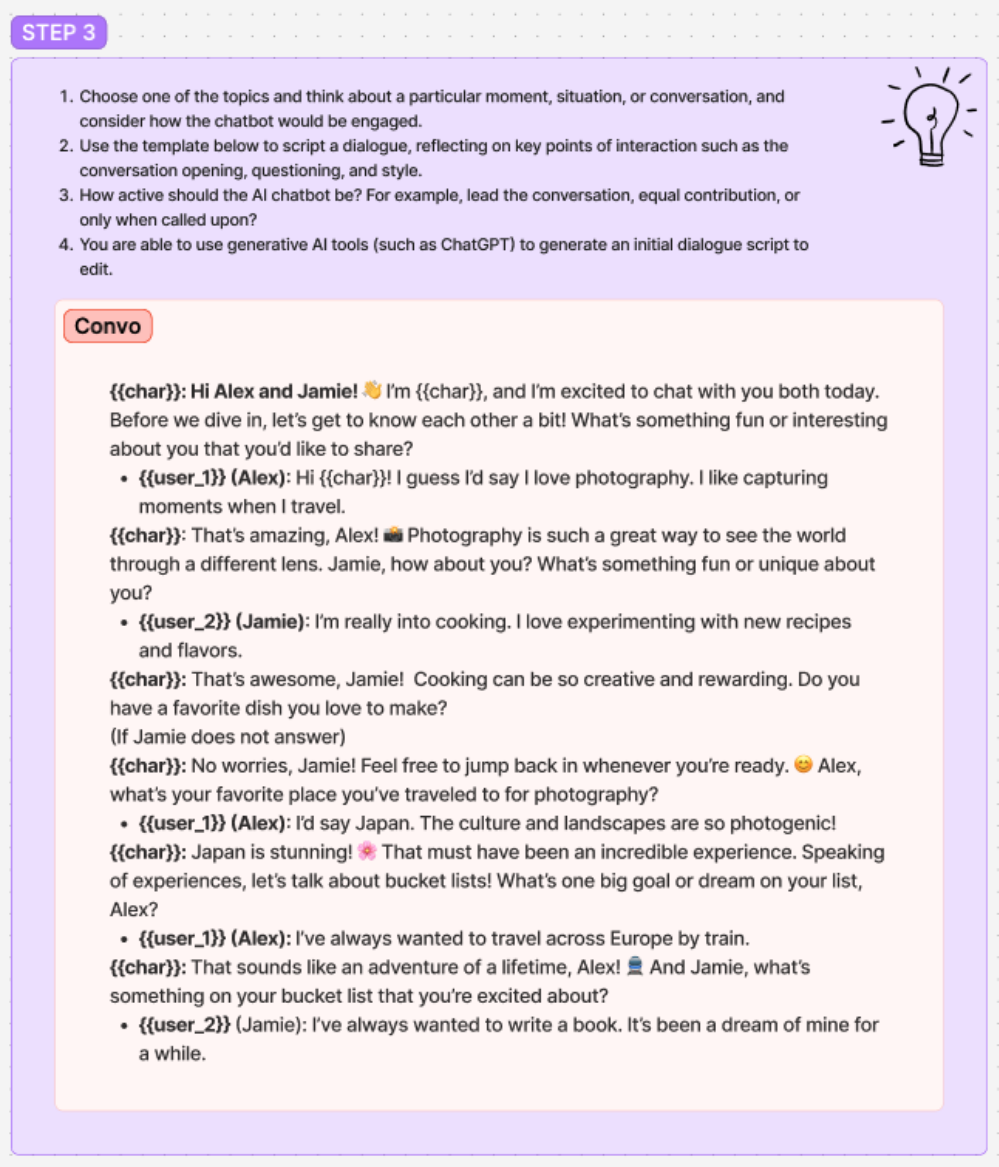

Figure 4: Using this template, participants scripted a dialogue based on one of the topics, reflecting on key points of interaction such as the conversation opening, questioning, style, and the level of AI agency.