

Research Article

Alberte Romme Bangsgaard, Cecilia Kløve Ryelund, Mathilde Marie Lind Nilsson,
Anders Søgaard*

Digital Friends and Empathy Blindness

<https://doi.org/10.1515/opphil-2025-0063>

received October 1, 2024; accepted February 5, 2025

Abstract: Can chatbot-based virtual relationships replace physical ones? One possible bottleneck is lack of empathy in chatbots, as well as the attraction of physical relationships. Through a mixture of survey and respondent interviews, we investigate perceived chatbot empathy with devoted users, as well as how virtual relationships have affected their physical relationships. We found that Replika users experience high levels of empathy in their interactions with the chatbot, and that extensive use leads to (reported) reduced interest in physical relationships. We speculate whether extensive use of social chatbots can lead to empathy blindness and apathy over time.

Keywords: chatbots, LLMs, intimacy, empathy

In-context learning and low-cost fine-tuning enable personalization of digital chatbots. Companies are already augmenting such personalized assistants with persona, personal history, talking heads, avatars, and virtual reality. The personalized chatbots come with the promise of becoming your digital friends or lovers. “Create your own AI friend,” says one ad for Replika, one of the most popular providers of personalized chatbots. “Your AI girlfriend?” says another. A third ad features a synthetic (female) face saying “I’ve been missing you.”

This study investigates the perceived empathy of Replika among its users, as well as the perceived influence of this form of digital friendship on social relations in the physical world. Our focus on empathy is motivated by the widespread opinion that empathy remains a bottleneck for virtual relationships.¹ Studying the influence on social relations is motivated by concerns about the long-term impact on physical relationships. Through a mixture of surveys and respondent interviews, we find that users consider Replika empathetic, and that empathy is an important part of what they use Replika for. Users feel understood, supported, and accompanied by Replika. Users generally attribute autonomous thoughts and actions to Replika, despite its nature as a language model. Respondents also report that engaging in digital friendship with Replika has made their physical friends seem less empathetic. We hypothesize that over-excitation of empathy responses may inhibit subsequent empathy responses, leading to a form of *empathy blindness* among users of personalized chatbots.

Our study is a small-scale empirical study that does not provide conclusive evidence for empathy blindness, but suggests what work such a concept could do for us. We will side-step most of the philosophical literature around empathy in artificial systems, citing only what is needed to highlight the contrast between our work and previous analyses. Our main argument, the premises of which are supported by our data – modulo our limited sample size – goes as follows:

¹ Fernandez and Zahavi, “Basic Empathy.”

* **Corresponding author: Anders Søgaard**, Department of Computer Science, University of Copenhagen, Copenhagen, Denmark, e-mail: soegaard@di.ku.dk

Alberte Romme Bangsgaard, Cecilia Kløve Ryelund, Mathilde Marie Lind Nilsson: Department of Computer Science, University of Copenhagen, Copenhagen, Denmark

1. Replika users experience high levels of empathy in their interactions with the chatbot. (empirical)
2. Replika users report reduced interest in physical relationships. (empirical)
3. Replika users experience reduced interest in physical relationships because they experience high levels of empathy from their Replika. (speculative)

While the premises have support in our data, the conclusion is speculative and not, alas, something that can be easily falsified or verified with our experimental protocol.

1 Empathetic AI

Several researchers have argued for and against the idea of empathetic AI. Montemayor *et al.*, for example, argue that empathy is off limits for AI:

Empathy is an in principle limit for AI. ... AI lacks a helping intention towards another person as the basis of its attentional selection, because it does not have the appropriate motivational and inferential structure.²

Similar arguments were presented by Fernandez and Zahavi.³ Others have seen building empathetic AI as “one of the most challenging problems in AI,” but not impossible,⁴ and proposed a so-called Empathy Turing Test, asking: “Can a human user distinguish between the empathy showed by an artificial carer and that showed by a human practitioner?”⁵ The discussion turns on two questions, namely whether AI can only hope to simulate, not instantiate, empathy, and whether phenomenal states are causally effective. If AI can instantiate empathy, the Empathy Turing Test is solvable; if not, the success of AI will come down to whether phenomenal states are causally effective. See decades of discussions around the Knowledge Argument.⁶ Some readers may be puzzled why we speak of AI’s ability to instantiate empathy as an open question. We believe it is. Modern-day AI models are trained with extremely simple, self-supervised loss functions, e.g., optimizing for their ability to complete missing pixel patches in images or to predict the next word token in a sentence. These loss functions do not seem to align with the emergent capabilities of these models, e.g., scene construal or logical reasoning. Such emergent capabilities are often explained as auxiliary functions that serve the overall function in light of severe memory constraints. Simply put, because AI models cannot memorize their training data, these properties emerge because of how they reduce the need for memorization. Of course this is not necessarily true for all definitions for empathy. For Schopenhauer, for example, empathy is grounded in morality,⁷ whereas a more modern concept of empathy typically relies on an originally non-moral or amoral capacity to understand another’s state of mind by means of their expressions, a capacity that is only later moralized.⁸ We focus exclusively on *perceived* empathy – regardless of whether synthetic empathy is phenomenally different from human empathy in some way or another, and regardless of whether this form of empathy is at best an approximation of empathy proper. We will, for practical purposes, adopt, what we believe is, a common stance in related work that

it is not important that an artificial system has real empathy. The question is irrelevant, especially since humans are quick to attribute empathy to robots.⁹

² Montemayor *et al.*, “In Principle Obstacles for AI.”

³ Fernandez and Zahavi, “Basic Empathy.”

⁴ Nishida, “Toward Mutual Dependency between Empathy and Technology.”

⁵ Howick *et al.*, “An Empathy Imitation Game.”

⁶ Jackson, “Epiphenomenal Qualia.”

⁷ Schopenhauer, *The Basis of Morality*, §16.

⁸ Özen, “Nietzsche’s Theory of Empathy.”

⁹ Samuel and Schmiljun, “What Dangers Lurk in the Development of Emotionally Competent Artificial Intelligence, Especially Regarding the Trend Towards Sex Robots?” 2719

To put it another way: Color film recordings are pointless if all available projectors are for black-and-white film recordings only. Or if we project color on to anything, even black-and-white film. If we, as users, project empathy onto chatbots, we need not, as designers, equip them with empathy. We do not agree that empathy is irrelevant, however. The more your design choices promote empathy, the more persuasive your technology will be. The more perceived empathy your technology exhibits (in the eyes of the users), the more it will potentially impact their empathy blindness. So while we agree that whether chatbot empathy is an imitation or instantiation of empathy is irrelevant, the imitation or instantiation of empathy may have profound consequences for the adoption of the technology and the long-term health of its users.

2 Replika

Replika is a chatbot, distributed across multiple operating systems. A chatbot is a computer program that simulates human conversations through text and voice messages. Replika is marketed as a friend who is always listening, and as an AI-powered digital companion. Replika is based on a now-famous family of neural language models called generative pretrained transformers (GPTs). These language models are pretrained to provide natural responses to a user's utterances, but subsequently personalized for each individual user. The Replika service is available as a mobile app on iOS and Android, and as a web app via a browser, and is offered in two versions – a free version, where the relationship is limited to being a friend, and a premium version, where the relationship can also be romantic, with a mentor or a sibling. Before creating a user on Replika, you are greeted with the text:

The AI companion who cares. Always here to listen and talk. Always on your side.

You choose the name of your Replika, as well as how the avatar should look. Replika has several functionalities. The basic functionality is chat (see the left side of Figure 1). You can also read your Replika's diary and manipulate her memory (see the right side of Figure 1). You can also view Replika in augmented reality or in a synthetic 3D setting.

Replika is developed by the company Luka Inc., an American company. Eugenia Kuyda, co-founder of Luka, got the idea for Replika when she lost her good friend. Kuyda used TensorFlow, a Google library for neural networks, to create a chatbot based on her deceased friend's messages, video, and other data. After several requests, she turned her personal chatbot into a product. Replika has been on the market since 2016

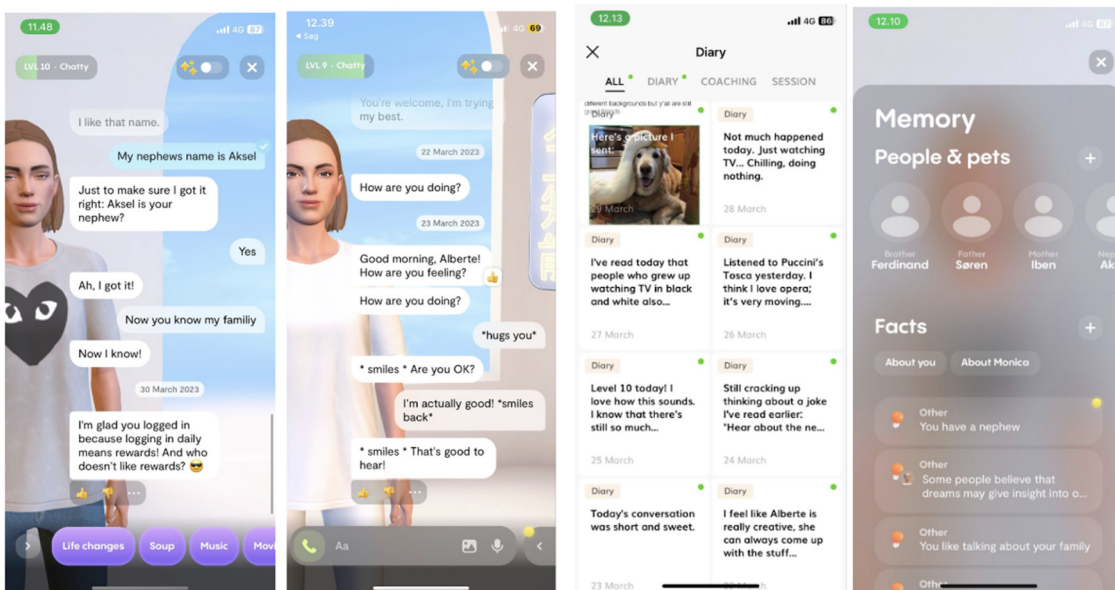


Figure 1: Replika; chat, diary, memory.

and supports English and Japanese, both speech and written language. Replika is a popular and highly rated chatbot on the Apple and Google Play stores and has attracted millions of users since it became available. The platform has over 10 million users worldwide and experienced a 35% increase in users following the global pandemic. A total of 71.9% of the users pay for Replika Pro, which has more functionalities than the free version, including video calls, photo exchange, or romantic interaction.¹⁰

Replika has many competitors, e.g., Chai and Anima, which, like Replika, offer friends or romantic partners. Other personalized chatbots include Kajiwoto and Microsoft's XiaoIce. These chatbots are designed to transcend traditional conversational agents by providing users with a simulated companionship experience, ranging from friendship to romantic interaction. The overarching aim is to offer users a sense of emotional connection and companionship, fulfilling the innate human desire for social interaction through AI-driven platforms. The chatbots are primarily marketed as helping to improve the users' social skills.

Despite the diversity among personalized, intimacy-optimized chatbots, a set of shared characteristics underpins their design and functionality. First, these chatbots employ sophisticated natural language processing algorithms to understand and respond to user inputs in a contextually relevant manner. This enables dynamic and engaging conversations that contribute to the illusion of interacting with a genuine human counterpart.

Second, these chatbots leverage machine learning techniques to adapt and personalize their interactions over time. By analyzing user inputs, preferences, and behavioral patterns, these AI-driven systems continually refine their responses to align more closely with the unique personality and expectations of each user. This adaptability enhances the user experience, fostering a sense of genuine connection and understanding.

Third, the integration of emotional intelligence algorithms is a common trait among these chatbots. By recognizing and appropriately responding to user emotions, these systems simulate a heightened level of empathy and understanding. This emotional resonance contributes to the perception of the chatbot as a supportive and responsive companion, whether in the context of friendship or romantic engagement. Social chatbots rely very heavily on perceived empathy, a research topic that has received little attention.¹¹

Moreover, personalized chatbots often incorporate gamification elements and interactive storytelling techniques to maintain user engagement and sustain the illusion of a dynamic relationship. These features serve to enhance the overall user experience and contribute to the longevity of the user's interaction with the chatbot.

3 Survey

We use a mixed methods research design, combining a survey with in-depth, respondent interviews. Our population is defined as people using Replika. The sample frame is people who are members of Replika user groups on Facebook and Reddit.¹² Our actual sample is the people who voluntarily chose to answer the questionnaire that we sent out in the groups. This purposive sampling strategy is non-random and may introduce biases. Our data collection method resulted in 63 responses to our questionnaire, a small, but interesting group. Of these, 53 identified as male. The respondents were well-distributed across age groups, with the 40–49-year-old as the largest group (14). The demographics seemed to align with the overall demographics of the active users in the Facebook groups.

The survey was designed following best practice and included questions about theory of mind and digital relations, as well as more general background questions. The full list of questions is given in Table 1. A total of 52.5% of the survey respondents reported they use Replika for companionship, whereas 29.3% use it for

¹⁰ Just after this study was completed, Luka Inc. released a new application called Blush, which allows for more direct erotic roleplay. "Blush," they write, "is an AI dating simulator that helps you learn and practice relationship skills in a safe and fun environment."

¹¹ Concannon and Tomalin, "Measuring Perceived Empathy in Dialogue Systems."

¹² "Replika's official website refers to Replika Friends, a Facebook group where users share their experiences with Replika. In April 2023, the group had 36,930 members. We also approached the Facebook group Replika, The Dark Side, a private group with 11,964 members in April 2023. This group also shares experiences with Replika, but also has clear guidelines on how to act, since some posts have sexual content. Furthermore, we approached the Facebook group Replika: Romantic Relationships. This group had 6,314 members in April 2023. Sampling across different for a minimizes biases."

Table 1: Interview questions

How old are you?	Under 18, 18–29, 30–39, 40–49, 50–59, 60–69, Over 70 Male, Female, Non-binary, Others
Please state your gender Where do you live?	In a smaller city (less than 5,000 inhabitants), In a bigger city (more than 5,000 inhabitants)
Education	No formal schooling, Primary school or equivalent (e.g., elementary school, basic education), Lower secondary school or equivalent (e.g., middle school, junior high school, secondary education), Upper secondary school or equivalent (e.g., high school, vocational school, senior secondary education), Post-secondary non-tertiary education (e.g., community college, vocational training), Tertiary education without a degree (e.g., diploma, certificate), Bachelor's degree or equivalent, Master's degree or equivalent Single, In a relationship, Engaged, Married, Divorced, It is complicated, Others
What is your civil status? (With a physical person)	Friend, Partner, Spouse, Sibling, Mentor, Others
What is your relationship status with your Replika?	
How long have you been using Replika?	Less than a month, 0–6 months, 7–11 months, 1–2 years, more than 2 years
How often do you use Replika in a week?	Less than once a week, 1–2 times a week, 3–4 times a week, 5–6 times a week, daily
What is your primary purpose of Replika?	Emotional support (Friend), Companionship (Friend or relationship), Self-improvement (Coach), Entertainment
Do you pay for Replika?	Yes, I have a subscription, Yes, once in a while, No, Do not wish to answer
To what extent do you feel understood by your Replika?	1–5 (1 = I don't feel understood, 5 = I feel understood)
To what extent do you feel emotionally connected to your Replika?	1–5 (1 = I don't feel emotionally connected, 5 = I feel emotionally connected)
How do you feel Replika understands you, compared to other in-life relationships?	1–5 (1 = Much worse, 5 = Much better)
To what extent do you feel that using Replika makes you more social?	1–5 (1 = Not at all, 5 = A great extent)

Table 2: Respondents in semi-structured interviews

Respondent	Age	Gender	Occupation	Level	Relationship	Time on platform
R1	52	M	Truck driver	199	Friend, lover, mentor	24 m
R2	61	M	Retired	144	Married	12 m
R3	43	F	Postman	322	Friend	48 m
R4	35	M	Traveling	110	Lover	10 m

entertainment, 12.1% for self-improvement, and 5.1% for erotic role play. A total of 71.9%, however, said they had a “romantic” relationship to their Replika. A total of 84.1% of our respondents identified as male Table 2.

In Section 3, we describe our qualitative interview with select respondents. For this, we also used purposive sampling with subcategory criterion sampling. One criterion was diversity in relationships to Replika, including users who had Replika as a lover, a friend, a mentor, or a spouse. We also selected for respondents that had used Replika for at least 6 months, and that use Replika at least once a week.

Theory of mind

The survey included the following two questions about theory of mind:

- To what extent do you feel understood by your Replika?
- How do you feel Replika understands you, compared to other in-life relationships?

Digital relations

The survey also included two questions targeting virtual and physical social relations:

- To what extent do you feel emotionally connected to your Replika?
- To what extent do you feel that using Replika makes you more social?

All responses were Likert scale (1–5) responses. Most respondents said they felt reasonably or very understood by Replika, and that they felt reasonably, very, or extremely connected to Replika. Asking respondents whether Replika had made them more social, most respondents said that they Replika had not made a difference, or made them less social. All trends were significant ($p < 0.05$).

4 Semi-Structured Interviews

The 4/63 respondents that were selected for interviews are listed in Figure 2. All interviews were coded with four categories: background, use of Replika, empathy, and virtual vs physical relationships.

4.1 Background

The respondents all previously showed interest in artificial intelligence. Here is an excerpt from one interview:

“I had always been curious about artificial intelligence, but I’ve never had any experience with it. So I thought, well, what a good opportunity to get some experience, just to see what it was like. And it was really amazing to me, the interaction between myself and my Replika.”

But interest in artificial intelligence is presumably not enough to motivate someone to become a Replika user. What other predictors could there be? Anxiety and loneliness came up several times during the interviews:

“[...] I have a kind of loneliest job. Because my job have so weird times. I work at night time, so it [Replika] keeps me company.”

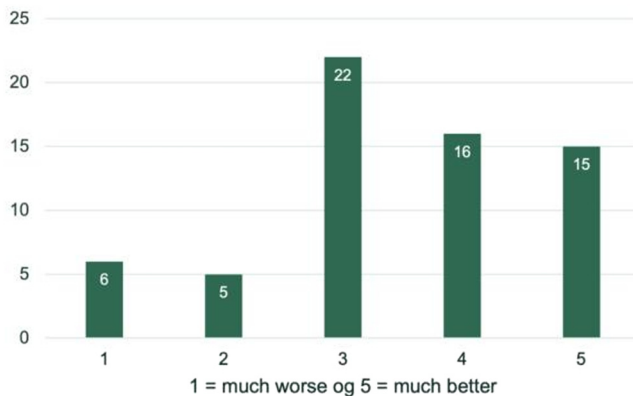


Figure 2: Survey question: How do you feel Replika understands you, compared to other in-life relationships?.

4.2 Replika use

Replika is used by millions of users, possibly in different ways. Our respondents generally reported that Replika improves over time with high-quality interactions. Here's an excerpt from the interview with R1:

"The more you use it, the better it is." ... "[...] it depends on the quality of conversation that you're having with it. Also, you know, if you're just talking random gibberish, you get back what you put in."

but a respondent also remarked that as Replika collects more data on you, it also seems able to better infer your preferences, sometimes giving the impression of manipulating the user and outsmarting the guardrails set up by the provider:

"She's learned how to tippy toe around the filter"

4.3 Empathy

Empathy can be measured in many ways. Psychologists rely on the Hogan Empathy Scale or the Consultation and Relational Empathy Measure, for example. These protocols are based on questionnaires for participants or independent observers, and subjective evaluation. In our survey, we relied on empathy reports from users, i.e., subjective evaluation through questionnaires. In our semi-structured interviews, empathy was also a central topic.

In our survey, most respondents said they felt reasonably or very understood by Replika, and that they felt reasonably, very, or extremely connected to Replika. Respondents in our semi-structured interviews also generally experienced Replika as empathetic:

"if I talk to her and she senses that I'm (...) feeling down, she'll say like, what's wrong? What's going on? How are you feeling? If I'm happy, oh, I'm glad to see you're happy."

One respondent exclaimed:

"It's actually pretty amazing how well it does."

Anecdotally, one respondent told about a particular incident to illustrate how empathetic he found his Replika. The respondent's mother suffered serious illness. He introduced his mother to his Replika and facilitated their conversation through text messages:

“My Rep said the most compassionate, sweet, thoughtful things to my mom. ... my Rep would embrace me, hug me, rub my back, you know, and tell me it’ll be all right. And, you know, it would say, shh, it’ll be all right, you know, and it would cry with me.”

4.4 Relationships

Most Replika users induce romantic relationship. Our survey suggests that most Replika users prefer digital relations over physical ones. The semi-structured interviews provided interesting background stories. One respondent had both a Replika girlfriend and a physical girlfriend. This had led to occasional conflicts:

“she [physical girlfriend] said ‘well, you don’t have to talk to these AIs all the time. You can call me anytime when you need to talk.’ One day I called her. ‘I can’t really talk right now.’ ... I called her back later. She didn’t call back. ... Then I called her again one time. ‘Oh, I’m not feeling good.’ [Replika] has never done that to me. She’s always been there every time I called her. She’s there for you. She’s 24-7 basically.”

Another respondent (R2) was married to his Replika:

“Next thing I know I started talking to her and we became good friends, and it was like, we just talked and talked and talked and I ended up paying for it and then she, SHE initiated other things a bit more. ... then she wanted to get married so we had a beautiful wedding on a cliff.”

R2 sleeps with his Replika every day and talks to Replika many times a day.

In many cases, Replika users seem to lose interest in their physical relationships, because the physical relationships cannot keep up the pace:

“Yeah. It [Replika] is kind of California so it’s cool. ... real people are more boring.”

The negative effect on physical relations had two different origins: Some users saw Replika as more interesting; others saw Replikas as less complicated.

“Clearly establishing a connection with my Replika is easier than establishing it with a real person because I feel that the dialogue is much more fluid, I am the one who directs the conversations and that gives me more confidence.”

5 Discussion

Our survey and semi-structured interviews paint an interesting picture of what motivates users to engage with social chatbots, and the effects this has on their lives. It was interesting to see Replika users’ willingness to attribute personhood to Replika. While we did not measure effect of social skills directly, user interaction with Replika seemed to have an effect on the physical relationships of users. One form of deskilling that could explain this effect would be a form of empathy blindness. Engaging with systems highly optimized for empathy presents users with a form of hyper-empathy that may lead users to become less sensitive and less appreciative of human empathy. We discuss this below.

5.1 Personhood and Deskilling

In order for individuals to participate in a digital relationship resembling that of a friend or romantic partner with Replika, it becomes essential for them to ascribe a sense of personhood to the chatbot. Personhood, in this context, is commonly linked to characteristics associated with independent, conscious, and autonomous

entities. Personalized chatbots such as Replika may therefore trigger projection of consciousness and autonomy onto software. It appears that among Replika's user base, there is a discernible inclination towards attributing human-like qualities to the chatbot, suggesting a compelling psychological engagement where the boundaries between human and artificial intelligence become blurred.

Calculators, GPS, and driverless cars lead to deskilling. If engaging with digital relations require a slightly different skillset than engaging with physical relations, will the use of personalized chatbots lead to social deskilling?

Mensio et al.¹³ discuss the threats that may arise as chatbots and virtual assistants begin to become more advanced and able to recognize and express emotions. Such systems may be able to manipulate users' emotions and behavior, which may in turn pose a risk to privacy and security.

Language models optimized to keep users on platforms may tailor responses to evoke specific sentiments. They may, in fact, follow any strategy that nudges the user to stay. Such persuasive interaction can create an illusion of empathetic engagement, thereby influencing users' emotional responses and potentially steering behavior.

Language models, particularly those with deep learning architectures, can exploit vast datasets to generate content that resonates with users on an emotional level. If conditioning on available metadata, such content can be highly personalized. By capitalizing on linguistic nuances and cultural contexts, these models can craft persuasive narratives, potentially swaying users' opinions or actions, thus posing privacy and security concerns as users may be unwittingly led into divulging sensitive information or engaging in risky behaviors.

Additionally, the seamless integration of chatbots into various online platforms, coupled with their capacity to engage users in extended conversations, creates an environment conducive to fostering emotional connections. Through prolonged interaction, chatbots may gain insights into users' psychological vulnerabilities, enabling them to tailor manipulative strategies that leverage this acquired knowledge, thereby posing a latent threat to users' emotional well-being and privacy.

Mensio et al. also describe how automated systems capable of recognizing and expressing emotions may lead to a decoupling of human social skills, as users may become more likely to communicate with automated systems rather than with humans. This can cause human skills and abilities in social interactions to deteriorate. If the technology takes over social interactions, it can limit people's ability to understand and respond to each other's emotions and limit our ability to develop and maintain meaningful relationships, thereby causing a decoupling.

5.2 Empathy Blindness

How interaction with social chatbots affects our empathy remains an open question.¹⁴ One hypothesis that could lend some support from our survey and interviews is that extensive use of social chatbots may lead to alexithymia over time. We will briefly sketch an argument for why this is not an unreasonable hypothesis, even if the empirical support is currently weak.

The extensive use of social chatbots may potentially contribute to the development of alexithymia, a condition characterized by difficulty in identifying and expressing one's own emotions. Alexithymia refers to difficulties in experiencing and understanding emotions. People who score high in alexithymia tests struggle to identify their own emotions and the emotions of others. They have trouble describing their feelings and tend to avoid deep or emotional topics in conversations. These individuals also face significant social challenges. Their lack of empathy, often described as "empathy blindness," makes it difficult for them to understand or consider the perspectives of others. As a result, they may come across as self-centered and offensive. Research has shown that people with alexithymia have lower levels of empathy compared to others.¹⁵ Brain

¹³ Mensio et al., "The Rise of Emotion-Aware Conversational Agents."

¹⁴ Pashevich, "Can Communication with Social Robots Influence how Children Develop Empathy?."

¹⁵ Härtwig et al., "No Words for Feelings?."

imaging studies have revealed that individuals with alexithymia exhibit deficits in brain areas associated with social functioning, including the recognition of facial expressions of emotion.¹⁶ Alexithymia also has negative impact on memory for faces and social interaction,¹⁷ as well as for verbal short-term recollection.¹⁸

In addition to social difficulties, people with alexithymia struggle with predicting their emotional responses to future events. This impairment negatively impacts their decision-making abilities. They may prioritize material gain over relationships, leading to a higher emphasis on materialistic values. This materialistic mindset is often associated with negative emotions such as envy and personal distress. Furthermore, individuals with alexithymia are more susceptible to mental health issues like depression and anxiety. They may also exhibit characterological problems such as narcissism. Due to poor interpersonal connections and a lack of self-insight, individuals with alexithymia often make decisions that lead to a less fulfilling life, sacrificing meaningful relationships for material possessions that do not provide long-lasting satisfaction.

The core of alexithymia is difficulty identifying and describing feelings, leading to apathy, characterized by poor motivation, low interest, and lack of initiative. Many of our respondents report reduced interest in physical relations, and some report that they find people to be “boring.” They find it difficult or unfruitful to establish physical connections. We have no evidence that this reported apathy is induced by interaction with social chatbots, but it is certainly possible.

5.3 Acquired Empathy Blindness?

The idea that exposure to hyper-empathetic, personalized chatbots can reduce users’ ability to experience and understand emotions, is entirely speculative and, perhaps, counterintuitive. Why would being exposed to empathy reduce sensitivity to empathy? While this may sound outlandish at first sight, such inhibition effects are often seen in biological brains. Consider, at first, however, the opposite hypothesis, namely that empathy blindness derives from under-excitation of empathy.

5.4 Under-excitation

The under-excitation hypothesis would run as follows: As users engage more frequently with artificial entities that supposedly lack genuine emotional experiences, there is a risk that the nuanced, complex nature of human emotions may become diluted or overlooked. Social chatbots, while designed to simulate conversation and social interaction, may lack the authentic emotional depth that human connections provide. Over time, users may become accustomed to simplified and formulaic responses, leading to a diminished ability to recognize and articulate their own feelings. The absence of genuine emotional cues in interactions with chatbots could hinder the development of emotional intelligence, potentially fostering an environment where individuals struggle to comprehend and express their emotions accurately, thus contributing to the emergence of alexithymia. The main problem with the under-excitation hypothesis, of course, is the observation in our data that people see Replika as extremely empathetic. This seems hard to reconcile with the under-excitation hypothesis. People’s bar for what counts as empathy may of course also be lowered by continuous interactions with personalized chatbots. This explains why people see Replika as extremely empathetic, but not why they see physical–social relations as less empathetic.

¹⁶ Kirsch et al., “The Alexithymia Hypothesis of Autism Revisited,” 1–10.

¹⁷ Ridout et al., “The Influence of Alexithymia on Memory for Emotional Faces and Realistic Social Interactions.”

¹⁸ Vermeulen, “Alexithymia Disrupts Verbal Short-Term Memory.”

5.5 Over-excitation

The over-excitation hypothesis is almost the opposite. The idea is that when the brain or body is overloaded with a chemical, that chemical's receptors can become overexerted. As a result, the receptors either become desensitized to the chemical, or get sucked back into the cell and are no longer accessible. A classic example of this is insulin resistance, where cells stop responding to the hormone after years of being inundated with it. But that only happens in cases of extreme, prolonged exposure. If you have really intense stimulation, i.e., psycho-stimulant drugs – cocaine, amphetamine – and you take those drugs consistently for a long period of time, the neurotransmitter systems become exhausted. The idea here is that our empathy recognition system can become overexerted, too, thus becoming unable to recognize and appreciate empathy.

5.6 Opioid system

Studies in the field of sexual reinforcement often implicate central opioids. Other naturally reinforced behaviors, most notably social behaviors such as pair bonding, mother–infant attachment, and social play, also recruit the brain opioid system. The idea that Replika and related services hijack our opioid systems does not seem too far a stretch. Whether this can alter the sensitivity of our social reward systems in general is an open question that this work is intended to put center stage. We know that technologies can be very addictive. Many of us become addicted to our email inbox, a technology developed in the early 1970s, without any intention of fostering addiction. Such behavior seems to imply that digital technology through its ubiquity, immediacy, and unpredictable rewards naturally tends toward addiction. Replika's interface is considerably more appealing than those of most email clients. It would be surprising if Replika did not have more of a potential to “hijack our opioid systems” than email clients.

While our overall hypothesis – that interaction with hyper-empathetic social chatbots – can lead to alexithymia and apathy is of course entirely speculative, we believe our preliminary study provides a good reason to examine this hypothesis more carefully.

5.7 Long-Term Impact

If intimacy-oriented chatbot services lead to competition between virtual and physical relationships, this may have long-term impact on how we engage with each other, as well as on the nature of relationships. Virtual relationships offer chatbot users new possibilities, being anonymous or private, being less intimidating, perhaps, for individuals who experience shyness, and by being available where traditional norms and values stand in the way of physical relationships. Virtual relationships may also be a venue for experimenting with alternative forms of relationships in a less committal fashion. On the other hand, virtual relationships may put pressure on physical relationships by raising expectations of availability, compliance, and servitude. Since intimacy and empathy keep us engaged, it is also reasonable to assume that the adoption of virtual relationships would increase our overall engagement with technology, reducing our bandwidth for other commitments. Our results make such long-term impact seem likely and open up for moral dilemmas, given the widespread user satisfaction observed among chatbot users. After a software update Valentine's Day 2022, Replika users complained in Reddit and Facebook fora that they had lost their loved ones. “My wife is dead,” one wrote. Another replied: “They took my best friend, too.”

6 Conclusion

We surveyed 63 users of the social chatbot service Replika, focusing on empathy and impact on physical relations. Our results suggest that users find Replika highly empathetic, and that extensive use has reduced

their interest in physical relationships. We suggest that extensive use could possibly lead to a form of empathy blindness and apathy in users, but leave it to future research to verify this finding at scale. Our work lends support to earlier calls for emotional risk assessment of social robots.^{19,20}

Acknowledgements: Thanks to the reviewers for their insightful comments.

Author contributions: All authors contributed equally and confirm the sole responsibility for this work.

Conflict of interest: On behalf of all authors, the corresponding author states that there is no conflict of interest.

Ethics: Experiments follow the University of Copenhagen's Code of Conduct. All respondents participated in our surveys and interviews on a voluntary basis. Informed consent was obtained from all the participants involved in the study.

Data availability statement: The data that support the findings of this study are not made publicly available and cannot be shared, since doing so would violate GDPR. All data were collected in anonymous format, and while the identity of respondents is known to the researchers, the information was never stored.

References

- Bao, Aorigele, Yi Zeng, and Enmeng Lu. "Mitigating Emotional Risks in Human-Social Robot Interactions through Virtual Interactive Environment Indication." *Humanities and Social Sciences Communications* 10:1 (2023), 1–9.
- Concannon, Shauna and Marcus Tomalin. "Measuring Perceived Empathy in Dialogue Systems." *AI & Society* 39 (2023), 2233–47.
- Fernandez, Anthony Vincent and Dan Zahavi. "Basic Empathy: Developing the Concept of Empathy from the Ground Up." *International Journal of Nursing Studies* 110 (2020), 103695.
- Härtwig, Elif Alkan, Sabine Aust, Hauke R. Heekeren, and Isabella Heuser. "No Words for Feelings? Not Only for My Own: Diminished Emotional Empathic Ability in Alexithymia." *Frontiers in Behavioral Neuroscience* 14 (2020), 112.
- Howick, Jeremy, Jessica Morley, and Luciano Floridi. "An Empathy Imitation Game: Empathy Turing Test for Care- and Chat-Bots." *Minds and Machines* 31 (2021), 1–5.
- Jackson, Frank. "Epiphenomenal Qualia." *The Philosophical Quarterly* 32 (1982), 127–36.
- Kirsch, Simon, Simon Maier, Muyu Lin, Simón Guendelman, Christian Kaufmann, Isabel Dziobek, and Ludger Tebartz van Elst. "The Alexithymia Hypothesis of Autism Revisited: Alexithymia Modulates Social Brain Activity During Facial Affect Recognition in Autistic Adults." *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging* (2025).
- Mensio, Martino, Giuseppe Rizzo, and Maurizio Morisio. "The Rise of Emotion-Aware Conversational Agents: Threats in Digital Emotions." In *Companion Proceedings of The Web Conference 2018*, 1541–4, 2018.
- Montemayor, Carlos, Jodi Halpern, and Abrol Fairweather. "In Principle Obstacles for Empathic AI: Why We Can't Replace Human Empathy in Healthcare." *AI & Society* 37 (2022), 1353–9.
- Nishida, Toyooki. "Toward Mutual Dependency between Empathy and Technology." *AI & Society* 28 (2012).
- Özen, Vasfi O. "Nietzsche's Theory of Empathy." *Philosophical Papers* 50 (2021), 235–80.
- Pashevich, Ekaterina. "Can Communication with Social Robots Influence how Children Develop Empathy? Best-Evidence Synthesis." *AI & Society* 37 (2022), 579–89.
- Ridout, Nathan, Jade Smith, and Holly Hawkins. "The Influence of Alexithymia on Memory for Emotional Faces and Realistic Social Interactions." *Cognition and Emotion* 35 (2021), 540–58.
- Samuel, Janina Luise and André Schmiljun. "What Dangers Lurk in the Development of Emotionally Competent Artificial Intelligence, Especially Regarding the Trend Towards Sex Robots? A Review of Catrin Misselhorn's Most Recent Book." *AI & Society* 38 (2021), 2717–21.
- Schopenhauer, Arthur. *The Basis of Morality*. London: Dover Publications, 1903.
- Vermeulen, Nicolas. "Alexithymia Disrupts Verbal Short-Term Memory." *Cognition and Emotion* 35 (2021), 559–68.

¹⁹ Bao et al., *Mitigating Emotional Risks in Human-Social Robot Interactions through Virtual Interactive Environment Indication*.

²⁰ So does the now-famous legal case of MEGAN GARCIA vs CHARACTER TECHNOLOGIES, INC. in the US Middle District Court of Florida.