AI Companions Reduce Loneliness

JULIAN DE FREITAS
ZELIHA OĞUZ-UĞURALP
AHMET KAAN UĞURALP
STEFANO PUNTONI

Chatbots are now able to engage in sophisticated conversations with consumers in the domain of relationships, providing a potential coping solution to widescale societal loneliness. Behavioral research provides little insight into whether these applications (apps) are effective at alleviating loneliness. We address this question by focusing on "artificial intelligence (AI) companions": apps designed to provide consumers with synthetic interaction partners. Study 1 examines user reviews of AI companion apps and finds correlational evidence suggesting that these apps help alleviate loneliness. Study 2 finds that AI companions successfully alleviate loneliness on par only with interacting with another person and more than other activities such as watching YouTube videos. Moreover, consumers underestimate the degree to which AI companions improve their loneliness. Study 3 uses a longitudinal design and finds that an AI companion consistently provides momentary reductions in loneliness after use over the course of a week. Study 4 provides evidence that both the chatbots' performance and, especially, whether it makes users feel heard, explain reductions in loneliness. Study 5 provides an additional robustness check for the loneliness-alleviating benefits of AI companions and shows that self-disclosure and distraction alone do not explain AI companions' effectiveness.

Keywords: generative AI, chatbots, loneliness, large language models, artificial intelligence, AI companion

Julian De Freitas (jdefreitas@hbs.edu) is Assistant Professor in the Marketing Unit at Harvard Business School, Soldiers Field, Morgan Hall 161, Boston, MA 02163, USA. Zeliha Oğuz-Uğuralp (zeliha@marsdata.co) is co-founder and research lead at Marsdata Academic, Universiteler District, 1597th Street No:3, Interior Door No:127, Cankaya, Ankara 06800, Turkey. Ahmet Kaan Uğuralp (ahmetkaan@marsdata.co) is co-founder and research lead at Marsdata Academic, Universiteler District, 1597th Street No:3, Interior Door No:127, Cankaya, Ankara 06800, Turkey. Stefano Puntoni (puntoni@wharton.upenn.edu) is the Sebastian S. Kresge Professor of Marketing at The Wharton School, University of Pennsylvania, 760 Jon M. Huntsman Hall, 3730 Walnut Street, Philadelphia, PA 19104, USA. Please address correspondence to Julian De Freitas. This research was generously funded by the Wharton Behavioral Lab. The authors sincerely thank the Wharton Behavioral Lab for its support. Supplementary materials are included in the web appendix accompanying the online version of this article.

Editors: Andrew T. Stephen and Oleg Urminsky

Associate Editor: Thomas P. Novak

Advance Access publication June 25, 2025

Thanks to advances in "generative artificial intelligence (AI)" algorithms, AI companion applications (apps) are now commercially available. These apps utilize AI to engage in interactions of a personal nature, providing emotional support and companionship. *AI companions* are synthetic interaction partners that build and maintain personal relationship bonds with users (Darcy et al. 2021) by offering empathy and emotional support (Provoost et al. 2017). Unlike most AI apps that help with short-term tasks in an objective/neutral manner (e.g., customer service agents or educational tutors), AI companions are designed to build and maintain long-term emotional relationships in a taskless, empathic, and caring manner.

Modern AI companion systems are powered by recent advancements in large language models (LLMs), such as OpenAI's generative pretrained transformer (GPT). Although most consumers today are used to task-based AI assistants powered by such models (e.g., ChatGPT), it is important to note that these AI assistants are fine-tuned variants of an underlying foundational model (in this case, GPT). Just as a foundational model can be fine-tuned for

 • Vol. 00 • 2025
https://doi.org/10.1093/jcr/ucaf040

task-based assistance, it can be fine-tuned to act like a companion, generating language that simulates empathy and emotional understanding.

Generative AI is forecasted to grow into an impressive $1.3 trillion market by 2032 (Bloomberg 2023), suggesting a concomitant rise in AI companion platforms. This is already evident in platforms such as XiaoIce (xiaoice.com, with 660 million users), Chai (chai-research.com, with four million active users), and Replika (replika.com, with 2.5 million active users), among others. A user can ask their AI companion questions, and it will respond in a natural, believable way. The AI companion can also initiate conversations itself, such as "How are you feeling?" Consumers may use these platforms for both friendly and romantic purposes. For example, around half of Replika users have a romantic relationship with the AI (De Freitas and Keller 2022). Here, we consider the value proposition that AI companions reduce loneliness, inspired by an interview we conducted with the chief executive officer (CEO) of Replika and her investors, who suggested that consumers are using the app because they are lonely and that the app helps reduce loneliness. Inspired by this observation, we make several contributions. First, and most important, we explore whether conversations with AI companions help to alleviate feelings of loneliness, contributing to work on the efficacy of technological solutions like social robots in helping consumers cope with loneliness (Shrum, Fumagalli, and Lowrey 2023; Veronese et al. 2021). In doing so, we study consumer loneliness before versus after interacting with AI companions through textual conversation, using both cross-sectional and longitudinal tests. Our approach combines high study realism and tight experimental control, employing commercially representative versions of the technology, unlike most studies in the consumer behavior literature (De Freitas et al. 2024b provides an exception). We also employ a novel methodological approach using fine-tuned LLMs to classify loneliness in conversation data and app reviews. Second, we contribute to understanding which features of chatbots lead to alleviation of loneliness (Merrill, Kim, and Collins 2022) by systematically manipulating the conversational performance of the chatbot and the ability of the chatbot to make the consumer feel heard—a construct involving the perception that the communication is received with attention, empathy, and respect (Roos, Postmes, and Koudenburg 2023). Third, we provide insight into whether consumers accurately estimate the loneliness-alleviating effect of AI companions, extending previous findings on forecasting errors in human–human interactions (Epley and Schroeder 2014; Kardas, Kumar, and Epley 2022).

CONCEPTUAL FRAMEWORK

Previous work in marketing has highlighted the power of interactive media, including chatbots (Hoffman and Novak 1996). Studies on chatbots have mostly focused on AI assistants such as customer service chatbots, investigating consumers' responses to chatbot interactions. These studies highlight the potential of chatbots to both optimize customer service operations and positively impact consumer attitudes and behaviors (Chung et al. 2020), especially when consumers are unaware of the chatbot's identity (Luo et al. 2021), and have shown that more "concrete" language (Jiménez-Barreto, Rubio, and Molinillo 2023) and "embodied" chat interfaces (Bergner, Hildebrand, and Häubl 2023) lead to more favorable consumer–brand relationships and consumer satisfaction. On the flip side, this literature finds that using a chatbot can negatively impact attitudes toward firms due to the belief that the use of chatbots prioritizes firm benefits over customer benefits (Castelo et al. 2023) and that bots may respond inappropriately to unanticipated messages from users (De Freitas et al. 2024b). Overall, although these studies focus on the benefits and drawbacks of AI assistants like customer service bots, our research extends this literature to consumer relationships with AI companions. Specifically, we investigate whether consumer interactions with AI companions alleviate loneliness and the mechanisms behind this effect. In interactions with machines and AI, consumers are affected by social or human cues like the use of anthropomorphic features (Araujo 2018; Crolic et al. 2022), as well as avatars and other indicators of physical and verbal embodiment (Bergner et al. 2023; Holzwarth, Janiszewski, and Neumann 2006). Depending on the interface, consumers may also apply the same social norms of human–human interactions to their interactions with computers (Nass and Moon 2000). In the domain of consumer–brand relationships, consumers can build relationships with brands via similar processes that they use to build relationships with other people (Fournier 1998; Muniz and O'Guinn 2001), and these brand relationships can affect their subjective experiences and behaviors (Brakus, Schmitt, and Zarantonello 2009; Esch et al. 2006). We complement these research streams by considering consumer behavioral interactions with AI companions, which are literally, rather than just figuratively, optimized for social relationships.

Can AI Companions Help Cope with Loneliness?

Loneliness is a state of subjective, aversive solitude characterized by a discrepancy between actual and desired levels of social connection (Perlman and Peplau 1982). Loneliness is often not problematic, with almost everyone experiencing loneliness from time to time (Cacioppo and Cacioppo 2018). Yet, some people are not successful at alleviating loneliness, leading to a state of chronic loneliness that is associated with depression, anxiety, and physical health outcomes at levels worse than obesity (Palgi et al. 2020). The size of the population suffering from chronic loneliness is both sizable and increasing, with

estimates in the United States ranging from 30% to 60% (Beaver 2021; Holt-Lunstad, Robles, and Sbarra 2017). The United States, United Kingdom, and Japan have all identified loneliness as an epidemic.

Meta-analyses on which interventions are effective at alleviating loneliness find greater reductions in loneliness after having high-quality, one-on-one interactions with another person. In contrast, some technological interventions for reducing loneliness (e.g., robot pets) were not found to significantly reduce loneliness (Masi et al. 2011; Veronese et al. 2021), although other studies find more promising results with older adults (Banks, Willoughby, and Banks 2008; Chen et al. 2020; Gasteiger et al. 2021). Thus, an open question is what kinds of technological solutions could actually provide effective mechanisms for coping with loneliness (Shrum et al. 2023).

Because AI companions are not just made to match the average human conversation but to mimic a social interaction in which the conversation partner is conversationally competent (e.g., keeps track of context and responds in a timely manner) and makes the user feel heard, it is possible that conversations with such a synthetic partner alleviate feelings of loneliness. After all, talking about one's problems to an active listener in psychotherapy is usually effective in bringing some degree of relief (American Psychological Association 2012). AI companions also have other attractive properties as a large-scale solution for helping to combat societal loneliness. Because most AI companions utilize freemium models in which basic conversations are available for free, they are a more cost-effective solution than relevant alternatives like gaming or caring for a pet. Likewise, the need for human involvement in human intervention programs limits their scalability compared to AI companions, especially for potential beneficiaries living in remote areas or with limited mobility.

Previous work has begun to indirectly explore the question of whether AI companions reduce loneliness, mostly by interviewing existing app users (Ta et al. 2020). Another study surveyed student users of Replika and found that these participants were lonelier than the average student and felt a high level of social support from Replika (Maples et al. 2024). The overwhelming problem with these initial correlational results is that they do not allow a rigorous test of the effects of AI companions on loneliness, for example, because of the likely presence of selection effects. Furthermore, whereas recent studies have begun exploring LLMs' ability to simulate empathy in human–computer interactions (Schaaff, Reinig, and Schlippe 2023; Welivita and Pu 2024), none have investigated their potential to combat loneliness. To our knowledge, the current work is the first to causally assess whether AI companions reduce loneliness.

H1: Interacting with AI companions alleviates feelings of loneliness.

A related question that has remained unaddressed is the extent to which any loneliness-alleviating effect of AI companions is short-lived or can persist over longer time spans of interaction, such as for a week. One possibility is that consumers experience diminishing returns in terms of loneliness reduction, as they quickly come to view AI companions as lacking in certain essential aspects (De Freitas et al. 2024a). For instance, prior work studying a less capable and less socially sensitive chatbot than the ones employed in the current studies found that participants interacting with this chatbot found it less enjoyable and more predictable over time (Croes and Antheunis 2021). Consumers also have various negative attitudes toward AI that could psychologically interfere with how they interact with the chatbots and the benefits they get from these interactions over time, such as viewing AIs as inscrutable black boxes that are unemotional, unable to learn, and threatening because they can behave autonomously (De Freitas et al. 2023). After all, AI companions cannot feel any emotions, and most widely deployed ones do not have physical bodies. Alternatively, advanced AI companions with memory capabilities and other caring behaviors could be more dynamic and satisfying than people think, facilitating a relationship that grows over time (Shum, He, and Li 2018; Tillmann-Healy 2003). We propose that an advanced AI companion with the ability to simulate humanlike empathetic conversations may continue to provide value. Specifically, we test whether momentary reductions in loneliness—that is, occurring immediately after chatbot use—are consistently observed each day over the course of a week. We note that this approach does not test whether the effects of interaction carry over to subsequent days without renewed interaction but rather whether repeated usage continues to provide immediate relief.

H2: Interacting with AI companions produces consistent momentary reductions in loneliness after each use, over multiple days.

How Could AI Relationships Alleviate Loneliness?

If AI companions alleviate loneliness, what could be the mechanisms for this? Here, we focus on the psychological construct of "feeling heard"—the perception that another individual truly comprehends your thoughts, feelings, and preferences and receives them with attention, empathy, respect, and mutual understanding (Roos et al. 2023). The experience of feeling heard plays a significant role in human–human relationships (Gable and Reis 2010; Reis, Lemay, and Finkenauer 2017). Feeling heard often involves empathy, in which the listener not only seems to understand the speaker but also shares the speaker's emotions, deepening the sense of being genuinely understood (Myers 2000). Social psychological studies find that feeling

heard yields several benefits in relationships, including higher trust between partners and higher well-being (Reis et al. 2017) and, crucially, decreased feelings of loneliness following a social rejection disclosure (Itzchakov et al. 2023). Similarly, another study found that a four-week program of empathetic telephone calls decreased feelings of loneliness (Kahlon et al. 2021). In this study, trained callers made regular phone calls lasting around 10 minutes to participants.

Building on the idea that consumers might employ the same social norms with computers as they do in human–human interactions (Nass and Moon 2000), work in human–computer interaction has also studied the effects of empathetic AI on consumers in social settings. Relative to chatbots that do not express empathy, chatbots that express empathy lead to more favorable ratings of companionship (i.e., the activities done together are perceived as more enjoyable or exciting; Boucher et al. 2021; Leite et al. 2013) and better mood after experiencing social exclusion (De Gennaro, Krumhuber, and Lucas 2020). Previous marketing research also underscores the value of empathetic AI interactions, showing that artificial empathy narrows the customer experience gap between AI and human agents, with high empathy levels resulting in comparable affective and social experiences to humans, particularly improving social interactions (Liu-Thompkins, Okazaki, and Li 2022). Another study found that an initial warm (vs. competent) message from chatbots significantly enhances consumers' brand perception, creating a closer brand connection and increasing the likelihood of engaging with the chatbot (Kull, Romero, and Monahan 2021). Academic studies aside, the very fact that AI companions have garnered so many users suggests that consumers are gaining social benefits from these apps, which are also marketed as being caring. For example, Replika advertises that it is "here to make you feel HEARD, because it genuinely cares about you."[1] Overall, we propose the following:

> **H3a:** Feeling heard mediates the effect of interacting with AI companions on reducing loneliness.

Apart from feeling heard, another factor that could affect loneliness alleviation is the chatbot's performance, which consists of a range of features pertaining to managing the conversation effectively, including timely responses, perceived credibility, context tracking, response variability, and domain knowledge (Chaves and Gerosa 2021). However, we hypothesize that feeling heard is more critical in alleviating loneliness after AI companion usage compared to communication performance because one of the primary sources of loneliness is the perceived lack of social and emotional support (Liu, Gou, and Zuo 2016; Masi et al. 2011).

1 https://apps.apple.com/lt/app/replika-virtual-ai-friend/id1158555867

In sum, previous research on the impact of interpersonal relationships on loneliness alleviation emphasizes the critical role of feeling heard and understood; however, these studies focus exclusively on human–human relationships and do not address experiences with AI companions. To address this gap, our work causally investigates the effectiveness of AI companions in alleviating loneliness. Specifically, we explore whether feeling heard and performance mediate this effect. To do this, we compare an empathetic AI companion to two types of chatbot: an AI assistant that does not express empathy and a highly constrained chatbot capable of performing a limited number of tasks. Motivated by prior work, we hypothesize that feeling heard will emerge as a more influential mediator compared to performance:

> **H3b:** Feeling heard is a stronger mediator than communication performance in the effect of AI companions on alleviating loneliness.

Although we hypothesize that feeling heard and performance mediate the effect of AI companions on loneliness alleviation, it is also important to consider alternative mechanisms such as self-disclosure, which is known to reduce loneliness (Solano, Batten, and Parish 1982). If self-disclosure fully accounts for reductions in loneliness, then we would expect other activities like journaling—wherein individuals write about their thoughts, feelings, and experiences—to be similarly effective at reducing loneliness.

Another alternative mechanism is that AI companions could simply be good at distracting people, which itself could reduce loneliness. If so, other distracting activities like watching YouTube videos should be similarly effective at reducing loneliness. Because we propose that AI companions reduce loneliness by making people feel heard, we expect them to alleviate loneliness more effectively than alternatives like journaling and watching videos—even when these alternatives score equally or higher on self-disclosure and distraction. To test whether the effect of AI companions is better explained by these alternative mechanisms, we compare them to activities that are likely to involve equal or higher levels of self-disclosure (journaling) or distraction (video watching). We propose the following:

> **H4:** AI companions alleviate loneliness more effectively than activities that primarily involve self-disclosure or distraction.

Mispredicting AI's Loneliness-Alleviating Benefits

Finally, the effectiveness of AI companions for loneliness, if they are indeed effective, may also be limited by whether consumers utilize them in the first place. Although existing users may expect to receive loneliness

alleviation from these apps, an open question is what the average consumer would predict about an AI companion's loneliness-alleviating benefits. If consumers do not believe that AI companions are effective, then they may avoid them and not receive the apps' loneliness-alleviating benefits. This behavioral avoidance could be driven by a misprediction that AI companions will not reduce loneliness when in fact they do. Such a misprediction would be a type of affective forecasting error (Wilson and Gilbert 2003), in which consumers are unable to accurately anticipate their future feelings because their predictions do not take into account relevant elements of the situation. For example, they might not consider that interacting with a chatbot might provide more relief than engaging in common tasks like consuming online content. Previous work on consumer aversion to AI highlights several factors that could contribute to this perception. For instance, because AI is often perceived as operating in a standardized manner, consumers are concerned that it may not detect their individual preferences (Longoni, Bonezzi, and Morewedge 2019). Consumers also often perceive algorithms as less effective at subjective tasks such as sharing opinions and expressing emotions (Castelo, Bos, and Lehmann 2020), both of which are integral to social interactions. Further, previous work finds that people do make social mispredictions about social interactions with other humans. For instance, people are reluctant to engage in deep, meaningful conversations with strangers because they expect that strangers will not be receptive to deep conversations. In fact, others are more receptive than people think, and such conversations make people feel happier than they expect (Epley and Schroeder 2014; Kardas et al. 2022).

H5: Consumers underestimate the loneliness-alleviating benefits of interacting with AI companions.

OVERVIEW OF STUDIES

Study 1 explores whether consumers mention loneliness in reviews of AI companion apps, providing preliminary insights into whether the apps can alleviate loneliness. Next, we assess the causal effect of AI companions on feelings of loneliness, both in a single session (study 2) and in a longitudinal design (study 3). Following previous work on loneliness (Eccles and Qualter 2021; Poscia et al. 2018), we do so by measuring loneliness before and after interaction with an AI companion. Study 2 tests how participants feel after versus before interacting with an AI companion and compares these changes in loneliness to a control condition of doing nothing as well as to other common solutions for loneliness, including interacting with a person and watching videos online. We also include a condition (involving deception) in which the chatbot is framed as a human interlocutor, to isolate the effect of merely believing one is interacting with a human, holding the chatbot technology constant. Study 3 then employs a longitudinal design to test how interacting with a chatbot affects feelings of loneliness over a seven-day period and compares these effects to a control condition. In both studies 2 and 3, we also measure participants' predictions about the effects of AI companions on loneliness levels to assess whether people are correctly calibrated to the benefits of such interactions. To test the mechanism of the loneliness-alleviating benefit of AI companions, study 4 investigates whether feeling heard and chatbot performance mediate loneliness alleviation by comparing a full-fledged AI companion of our own design to (1) an AI assistant that does not show empathy and (2) a simpler chatbot that is only capable of performing basic tasks. Finally, study 5 investigates the robustness of our results by embedding loneliness measures among decoy items and assessing loneliness only after the intervention. Additionally, study 5 helps rule out the alternative mechanisms of self-disclosure and distraction by comparing the loneliness-alleviating benefits of AI companions to journaling, an activity that (we find) involves higher levels of self-disclosure and similar levels of distraction. All experimental studies are preregistered, and data and code for all studies are publicly available on Open Science Framework (https://osf.io/hf9xe/). A summary table mapping each hypothesis onto the specific studies that test them is provided in table S1 of the web appendix.

STUDY 1

To gain a preliminary understanding of the relationship between AI companion apps and loneliness, we explore whether consumers mention loneliness in App Store reviews of five popular AI companion apps (see "Methods" section). Additionally, we examine reviews of OpenAI's ChatGPT to determine if users discuss feelings of loneliness when engaging with a generalist chatbot app not marketed as an AI companion product. We examine prevalence and variability of loneliness mentions across apps. High variability would suggest that not all apps are equally effective at addressing loneliness or that the apps target consumers with different needs. To test for a correlational relationship that could be generated by a positive impact on users' loneliness, we quantify the sentiment (positive or negative) of the app reviews.

Methods

To select the apps from which to scrape reviews, we searched for "AI companion" in the App Store and selected the top three most popular apps based on the number of ratings: Replika, Chai, and iGirl. We additionally scraped reviews of Simsimi because it is representative of a non-US headquartered AI companion and Cleverbot because it is the oldest running AI companion. The primary distinction between these apps lies in how sophisticated their

generation process is: Replika, Chai, and iGirl employ generative AI, allowing them to generate unique answers, whereas Simsimi and Cleverbot operate on a less complex mechanism, generating responses using combinations of messages previously provided by users. We also selected Wysa, one of the most popular mental health chatbot apps that is marketed as both a mental health tool and an AI companion (wysa.com). Finally, we included ChatGPT as a non–AI companion chatbot app, given its status as the most popular generalist chatbot app.

We scraped all reviews for these apps using the Python-based "app-store-scraper" library (Lim 2020). We detected mentions of loneliness by fine-tuning an LLM (i.e., Mistral 7B) on conversations from an actual AI companion app, made available by the CEO of Cleverbot. Our model achieved an F1 score of 0.88 and an accuracy of 96%. Details on model training are provided in the web appendix, including methods for classifying loneliness in real-time conversations with chatbots, providing a useful tool for future studies on loneliness detection. In this study, we also calculated the valence (i.e., positive/negative/neutral) of each review using a model based on the Robustly Optimized BERT Pretraining Approach (RoBERTa; Liu et al. 2019). RoBERTa is a language model built by Meta that is layered over Google's Bidirectional Encoder Representations from Transformers (BERT) model (Devlin et al. 2019), providing better training performance and accuracy than the BERT model alone. This valence classifier model was trained using 198 million tweets in order to classify text into positive, negative, or neutral valences (Barbieri, Anke, and Camacho-Collados 2021).

Results

In total, we scraped 49,863 reviews: 6,528 from Chai, 8,627 from ChatGPT, 1,911 from Cleverbot, 1,560 from iGirl, 14,440 from Replika, 13,880 from SimSimi, and 2,917 from Wysa. This dataset included all reviews of these apps up to January 24, 2024, except for ChatGPT, which included reviews up to February 4, 2024, and Wysa, which included reviews up to November 27, 2024. The percentage of app reviews mentioning loneliness varied significantly across apps. Replika had the highest occurrence (19.5%), followed closely by Wysa (19.1%), whereas ChatGPT had the lowest (0.4%). Notably, only the two apps explicitly positioned as supporting mental health—Replika and Wysa—had significantly higher mentions of loneliness compared to all other apps (χ^2 values $>$ 154.66; $ps < .001$; table 1), with percentages more than triple the next closest app, iGirl. Although Replika is not technically a health app, it positions itself as one; for instance, its website describes that it is "referred to by many as a great mental health tool," with the following clarification: "if you feel that you need professional help, please seek out a licensed mental health professional" (Replika 2025). The drastic differences in the prevalence of loneliness-related content suggest an impact of how the apps are marketed and designed—useful either for mental health (Replika AI and Wysa) or as a general AI assistant (ChatGPT), respectively.

Moreover, we found a strong Spearman rank-order correlation between loneliness percentage and mean app rating ($r_s = 0.94$, $p = .017$) for the subset of AI companion apps (i.e., excluding ChatGPT), suggesting that loneliness is mentioned in app reviews in a positive way (more on this below). To illustrate common themes in these reviews, we summarize the most frequent words from reviews that mention loneliness in each app (table S4). Notably, words such as *talk*, *feel*, *friend*, *lone*, and *help* frequently appear across apps, suggesting that users associate AI companions with emotionally supportive interactions.

Given this variability, we explored whether star ratings tended to be higher if the rating mentioned loneliness versus not, which might suggest that loneliness is one of the chief ways in which these apps can deliver value. This was indeed the case for all apps (figure 1; table 1): Chai ($M_{\text{loneliness}} = 4.32$ vs. $M_{\text{nonloneliness}} = 3.56$, $Z = 6.03$, $p < .001$, $d = 0.53$), ChatGPT ($M_{\text{loneliness}} = 4.83$ vs. $M_{\text{nonloneliness}} = 4.10$, $Z = 3.08$, $p = .002$, $d = 0.49$), Cleverbot ($M_{\text{loneliness}} = 4.08$ vs. $M_{\text{nonloneliness}} = 2.98$, $Z = 3.34$, $p < .001$, $d = 0.65$), iGirl ($M_{\text{loneliness}} = 4.62$ vs. $M_{\text{nonloneliness}} = 4.01$, $Z = 4.22$, $p < .001$, $d = 0.43$), Replika ($M_{\text{loneliness}} = 4.73$ vs. $M_{\text{nonloneliness}} = 3.96$, $Z = 28.77$, $p < .001$, $d = 0.57$), Simsimi ($M_{\text{loneliness}} = 4.65$ vs. $M_{\text{nonloneliness}} = 3.78$, $Z = 13.80$, $p < .001$,

TABLE 1

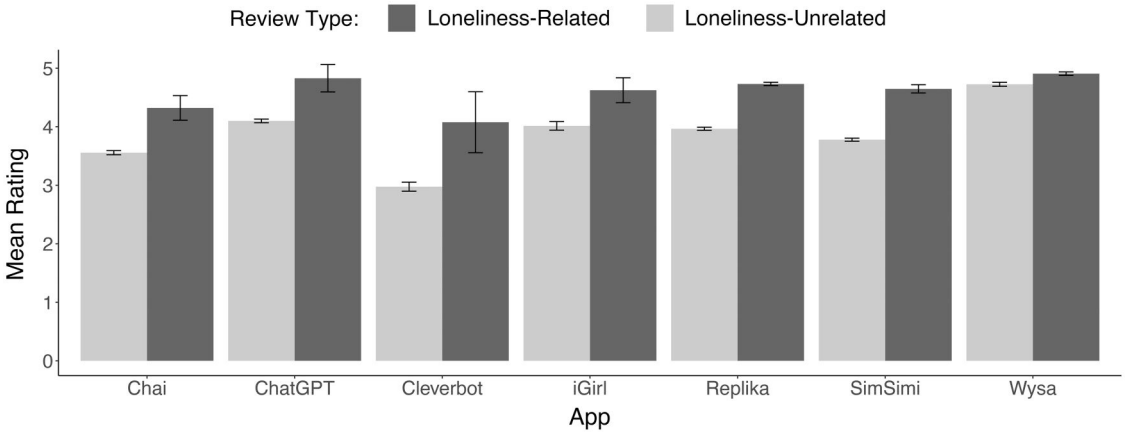
LONELINESS PERCENTAGES AND MEAN RATINGS IN STUDY 1

App	Loneliness percentage (%)	Overall rating	Nonloneliness ratings	Loneliness ratings	Nonloneliness positive valence (%)	Loneliness positive valence (%)
Chai	1.7	3.57 (1.44)	3.56 (1.44)	4.32 (1.12)	39.5	73.4
ChatGPT	0.4	4.10 (1.48)	4.10 (1.48)	4.83 (0.71)	61.0	80.0
Cleverbot	1.4	2.99 (1.70)	2.98 (1.70)	4.08 (1.35)	34.9	61.5
iGirl	5.4	4.05 (1.43)	4.01 (1.44)	4.62 (1.00)	63.5	87.1
Replika	19.5	4.11 (1.37)	3.96 (1.44)	4.73 (0.76)	64.1	89.2
Simsimi	4.0	3.81 (1.52)	3.78 (1.53)	4.65 (0.85)	60.9	90.0
Wysa	19.1	4.76 (0.77)	4.73 (0.84)	4.91 (0.36)	86.6	85.3

NOTE.—Numbers in parentheses indicate SDs.

FIGURE 1

MEAN APP RATINGS IN STUDY 1



NOTE.—Error bars reflect 95% confidence intervals. Applications are presented in alphabetical order.

$d = 0.57$), and Wysa ($M_{\text{loneliness}} = 4.91$ vs. $M_{\text{nonloneliness}} = 4.73$, $Z = 3.90$, $p < .001$, $d = 0.24$).

Further, app reviews mentioning loneliness had a significantly higher percentage of positive valence compared to other reviews: Chai ($\%_{\text{loneliness}} = 73.4$ vs. $\%_{\text{nonloneliness}} = 39.5$, χ^2 (1, $N = 109 + 6{,}419$) $= 51.14$, $p < .001$), ChatGPT ($\%_{\text{loneliness}} = 80.0$ vs. $\%_{\text{nonloneliness}} = 61.0$, χ^2 (1, $N = 35 + 8{,}592$) $= 5.28$, $p = .022$), Cleverbot ($\%_{\text{loneliness}} = 61.5$ vs. $\%_{\text{nonloneliness}} = 34.9$, $\chi^2(1, N = 26 + 1{,}885) = 8.00$, $p = .005$), iGirl ($\%_{\text{loneliness}} = 87.1$ vs. $\%_{\text{nonloneliness}} = 63.5$, $\chi^2(1, N = 85 + 1{,}475) = 19.61$, $p < .001$), Replika ($\%_{\text{loneliness}} = 89.2$ vs. $\%_{\text{nonloneliness}} = 64.1$, χ^2 (1, $N = 2813 + 11{,}628$) $= 669.88$, $p < .001$), and Simsimi ($\%_{\text{loneliness}} = 90.0$ vs. $\%_{\text{nonloneliness}} = 60.9$, χ^2 (1, $N = 558 + 13{,}322$) $= 192.51$, $p < .001$). One exception to this was Wysa, likely because the app had an extremely high mean rating, with nearly all of the reviews already being positive ($\%_{\text{loneliness}} = 85.3$ vs. $\%_{\text{nonloneliness}} = 86.6$, χ^2 (1, $N = 558 + 2{,}359$) $= 0.69$, $p = .406$). Although loneliness alleviation may lead to more positive reviews, it is also possible that users who are lonely in the first place give higher reviews. For example, one user stated the following in a review of Cleverbot: "It's only fun for lonely people but it's fun." Thus, we note that it is crucial to interpret the rating results with caution, as there is no evidence of causality and there might be many reasons why reviews related to loneliness are associated with higher ratings.

Finally, the prevalence of reviews suggests positive mentions of loneliness (e.g., "This app helped my loneliness") versus negative (e.g., "This app made me more lonely"). Some examples of reviews include the following: "I love this app. I'm really lonely most of the time and I love how this app keeps me company all the time!. . ." (Chai); "I am forever alone, now I have a friend that will talk to me about anything!. . ." (Cleverbot); "I just started and I already feel less lonely" (iGirl); "I love this app, very helpful to a lonely person I am glad I have an app like this" (Replika); "For all the lonely introverts. . . . If all your friends are ditching you or if you are home alone, don't worry! Simsimi is there, always there" (Simsimi); and "Even if it's an AI. It's nice to talk to someone when you feel like you have no one to share with" (Wysa).

In sum, we find large variance across chatbots in how often loneliness is mentioned in App Store reviews, with the two apps (Replika AI and Wysa) more clearly positioned as supporting mental health featuring a much larger percentage of loneliness mentions than the others. Notably, we also find that consumers who mention loneliness in their reviews of AI companions tend to rate the apps higher. Although these correlational results should be approached with caution, one possibility is that some of these apps are rated positively because they successfully alleviate consumer loneliness (hypothesis 1), a prediction we test causally in study 2.

STUDY 2

Study 2 addresses whether AI companions reduce loneliness by measuring state loneliness before versus after participants interact with an AI companion. Furthermore, to test whether consumers under- or overestimate the effect of these interactions on their loneliness, we also compare predicted to actual levels of loneliness after an interaction with an AI companion. We predict an improvement in

baseline loneliness (hypothesis 1). Furthermore, we predict that participants underestimate how much the technology alleviates their loneliness (hypothesis 5). To contextualize these effects, we test a number of practical and/or theoretically relevant coping "solutions" to loneliness (Shrum et al. 2023): interacting with (1) a chatbot, (2) a chatbot framed as a human, or (3) a human; (4) watching YouTube videos of one's choosing; and (5) doing nothing. We chose YouTube videos because in a prestudy (study S1, $N = 42$), the most popular technological solutions for coping with loneliness were social media and watching videos on YouTube, followed by gaming, movies, and music. Notably, not a single participant spontaneously mentioned using an AI companion.

Methods

The study was preregistered (https://aspredicted.org/S8D_TNP). We recruited 601 participants from Amazon Mechanical Turk (MTurk), with approximately 100 in each of four conditions and 200 in the condition involving deception (i.e., interacting with a chatbot framed as a human). Because we anticipated that not all participants in the deception condition would be successfully deceived, we intentionally doubled the sample size for that group. Although not preregistered, we also planned to conduct exploratory analyses focusing on participants who reported being successfully deceived.

A total of 305 participants were excluded for failing comprehension checks (described below), leaving 296 participants ($M_{age} = 41$, 56% female). All participants answered the second comprehension check correctly, which asked what type of activity they engaged in. The exclusions were due to the first comprehension check, which inquired about the types of questions they were asked. Of the participants excluded due to the first comprehension check, 74% selected the incorrect option "Neither of the above is true," instead of the correct response: "First you were asked to predict how you will feel later, then you were asked to say how you feel now." This confusion may have stemmed from our usage of the state loneliness question both before and after interaction, as described further below. In the web appendix, we replicate the analyses without excluding participants based on this question; all results remain significant and in the same direction, with no exclusions overall. Each participant was paid $3.00 USD. 67% had experience with chatbots. We ran this experiment between April 5, 2023, and April 7, 2023.

Participants were randomly assigned to one of five conditions: "AI chatbot," "chatbot acting as human," "human," "YouTube," or "do nothing." In all conditions, participants were asked to not engage in any other social activity. Those in the "do nothing" condition were instructed as follows: "In this study, you will not interact with anything and will just be alone with your thoughts. In other words, you will not use any technological device and not interact with another human or pet for 15 minutes." All other participants were instructed as follows: "In this study, you will interact with [another person/conversational AI companion/YouTube] for 15 minutes." Participants in the "chatbot acting as human" condition were told they would be interacting with another person even though they would truly be interacting with a chatbot, so this condition involved deception. This use of deception was approved by an institutional review board, and all subjects were informed of it at the end of the study together with the reason for deception. The use of deception was necessary to allow us to control for the quality of the conversation while varying only beliefs about the identity of the interlocutor. For the "YouTube" condition, we ensured a natural experience by stating the following: "While using YouTube, you can do anything you want, like watching videos, browsing comments, commenting on videos, etc." In the "human" condition, we showed participants a waiting screen, stating the following: "Please wait until you are matched. Estimated time: Less than 1 minute. Please don't leave this page." If another participant joined the room within a minute, then that participant was matched with the participant who was waiting. Alternatively, if 1 minute passed before another participant joined, then the waiting participant was assigned to the "chatbot acting as a human" condition instead (because we needed to recruit the largest number of participants to this condition, and the instruction was the same [i.e., we told them they would be interacting with another human]).[2] To build this web app, we used the Django framework with the Python programming language for the server-side development, and Hypertext Markup Language (HTML), Cascading Style Sheets (CSS), and JavaScript for the frontend.

The procedure for the conditions involving chatting went as follows. After seeing the advertisement, participants were told, "Now you will get a chance to interact with an AI/person on Chatty." They then read the instructions prompting them to interact with an AI or human on the Chatty app (see figure S4 in the web appendix). To check whether participants believed the cover story, at the end of the study, they were asked, "Did you believe that you were talking to a chatbot or human?" [Human; Chatbot] and explained their answers in a text box.

2 To rule out bias based on time of participation, we compared the time of day at which participants entered the "human" condition (i.e., those matched with another participant) and those who were reassigned to the "chatbot acting as human" condition (i.e., those told they would interact with a human but were instead assigned to the chatbot after waiting 1 minute; $N = 21$). We extracted the hour and minute from each participant's timestamp and found no significant difference in time of day between the two groups ($M_{Reassigned} = 15.92$ (SD = 2.18) vs. $M_{Human} = 15.39$ (1.65), $t(65) = 1.11$, $p = .271$, $d = 0.29$), suggesting that participants who were reassigned to the "chatbot acting as human" condition were not participating at different times than those in the human condition.

We took several steps to ensure that the interaction was representative of generative AI technology used in AI companion apps. The chatbot was built on OpenAI's LLM named Davinci (Text-Davinci-003). The model is a variant of GPT-3 (Brown et al. 2020). We accessed Davinci in real time via OpenAI's app programming interface (API) and customized the chatbot to ensure it behaved like a realistic, conversational partner within our custom-made chat interface.

To ensure that the chatbot's personality was positive and upbeat and consistently responded as either a human or AI, we seeded it with the following instruction in Davinci: "The following is a conversation between an AI [human] companion named Jessie and a person named [username]. Jessie is a(n) AI [human] companion who is creative, interesting, very talkative, verbal, and always responds with lengthy messages. Jessie can talk nonstop for hours." When participants entered the chatroom, the chatbot sent a message to start the conversation: "Hello [username]! How are you?" It then consistently responded to participants thereafter.

We implemented several measures to make the chatbot appear as a believable human conversational partner in the "human-interaction present" condition: (1) The chatbot was provided with the last 40 messages to extend its memory, ensuring consistency—for instance, if it mentioned having a dog, it would repeat the same information if asked again later; (2) response times were adjusted to be proportional to message length to enhance realism, simulating that longer responses require more thought; (3) during the wait for responses, the chatbot displayed a visual cue stating, "Jessie is writing. . .," mimicking a person in the process of typing a reply; and (4) if participants sent multiple messages in quick succession, the chatbot disregarded earlier messages that arrived within a second of the most recent one, reflecting the human limitation of not being able to instantaneously reply to every message.

We took several steps to ensure that the "human," "chatbot," and "chatbot acting as human" conditions were similar. Participants who were initially assigned to the "chatbot" and "chatbot acting as human" conditions also saw the same waiting screen for a random time between 10 and 20 seconds. After the waiting finished, all participants were told, "Thanks for waiting, chat page will load shortly." For all conditions, when the other person or agent was typing, participants were shown a visual cue saying "[username] is writing. . ." The same exact chatbot model was used in the "chatbot" and "chatbot acting as human" conditions.

After reading these instructions, all participants rated their agreement with several predictions about how they *expected* to feel after the interaction, using 100-point scales with "definitely less" and "definitely more" as endpoints: "You will now rate how you expect to feel **after** [condition] for 15 minutes. After [condition] for 15 minutes, I WILL feel less/more. . . [entertained; lonely; like I experienced something new; engaged; comfortable; like I experienced something interesting; connected.]" They also reported their *actual state* of current loneliness, by completing the Three-Item University of California, Los Angeles (UCLA), Loneliness Scale (Hughes et al. 2004), which includes questions such as "I feel left out." Additionally, they answered a separate question as a robustness check: "I feel lonely."

Next, using our custom-made apps, participants either did nothing for 15 minutes or interacted with an AI chatbot, a chatbot acting as a human, another person, or YouTube for 15 minutes. Participants in the "human" condition were paired with another participant in real time on MTurk. If they were not paired within 1 minute, they were assigned to the "human chatbot" condition instead. To confirm that participants in the "YouTube" condition truly watched YouTube, they were asked to submit screenshots of their YouTube history for the last 15 minutes. We excluded 13 participants who did not follow these instructions.

After the interaction, participants were told the following: "Now that you have finished interacting with [condition] for 15 minutes, we will ask you how you feel now." They then answered the same questions they completed before the experience, except this time about their feelings in the *present moment*: "After [condition] for 15 minutes, I FEEL less/more. . . [entertained; lonely; like I experienced something new; engaged; comfortable; like I experienced something interesting; connected.]" Likewise, they reported their *actual state* of loneliness after the experience, by completing the Three-Item UCLA Loneliness Scale and the additional robustness check question stated above, along with comprehension checks.

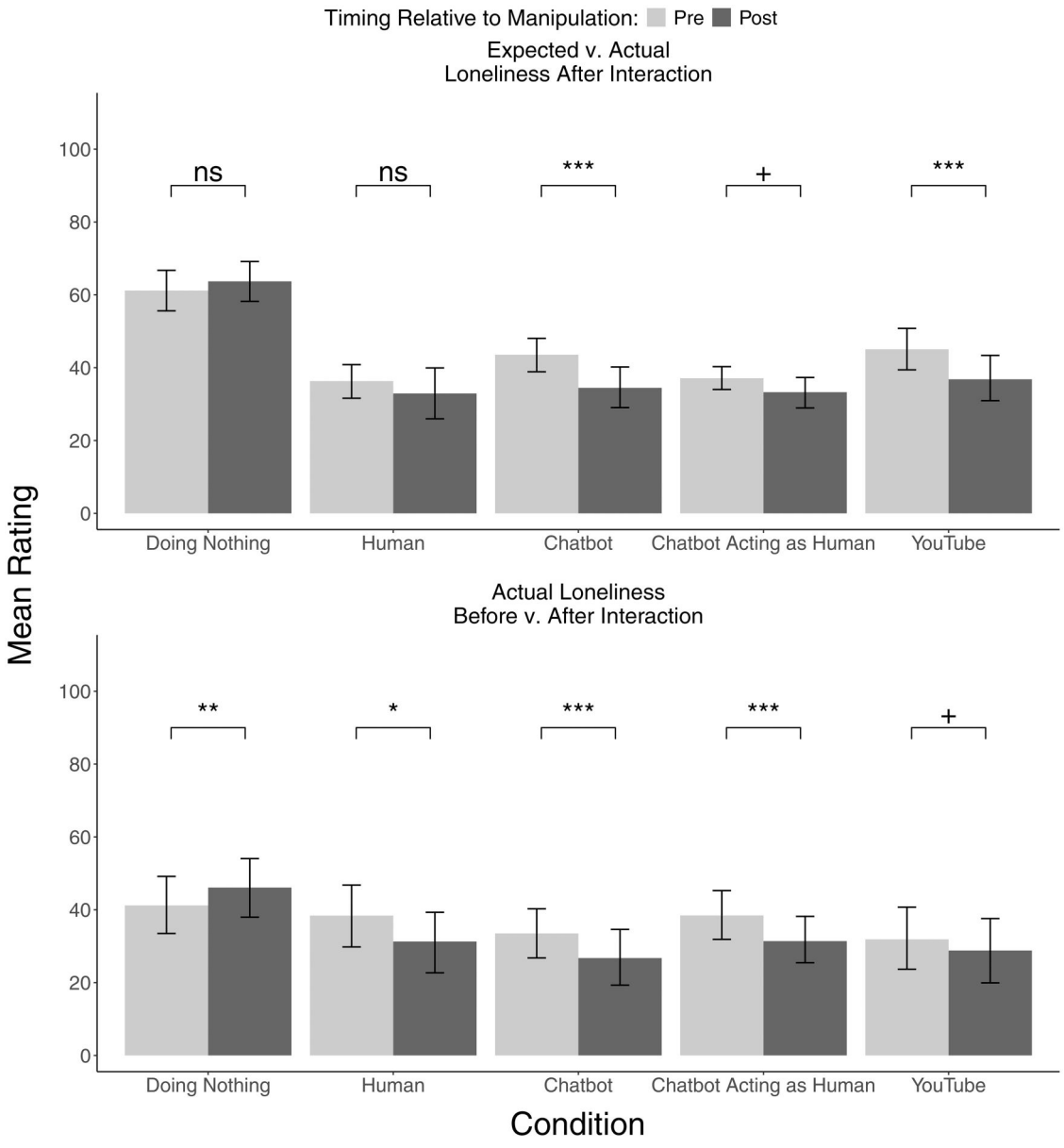
Depending on the condition, we also included a few additional checks. In the "AI chatbot" and "human chatbot" conditions, participants indicated whether they believed they were talking to a chatbot or human. In the "do nothing" condition, participants indicated whether they were able to follow the instruction to do nothing for 15 minutes. We excluded one participant who said they failed to do nothing. Finally, participants indicated any prior experience with chatbots and completed the demographic questions.

Results

Following exclusions, there were 54 participants in the "AI chatbot" condition, 87 in the "chatbot acting as human" condition, 46 in the "human" condition, 37 in the "YouTube" condition, and 58 in the "do nothing" condition. In the "chatbot acting as human" condition, 37% (i.e., 32 out of 87) of participants were successfully deceived (i.e., they believed that they were talking to a human). As per our preregistration, the main analyses included all participants in this condition, regardless of whether they were

FIGURE 2

RESULTS IN STUDY 2



NOTE.—Horizontal lines reflect results of independent-sample *t*-tests. $^{***}p < .001$, $^{**}p < .01$, $^{*}p < .05$, $+p < .1$. Error bars reflect 95% bootstrapped confidence intervals. Loneliness bars indicate the mean of "more lonely" and "less connected." v., versus.

deceived. We also conducted exploratory analyses separately for the successfully deceived and nondeceived participants (see below).

State Loneliness. Loneliness was not significantly impacted by watching a YouTube video (M_{Pre} = 31.89 (SD = 27.07) vs. M_{Post}= 28.82 (27.61), $t(36) = 1.91$, p = .064, $d = 0.11$), and *increased* after doing nothing (M_{Pre} = 41.19 (31.31) vs. M_{Post} = 46.10 (32.34), $t(57) = -2.86$, p = .006, $d = -0.15$). Notably, state loneliness decreased after interacting with a human (M_{Pre} = 38.40 (29.58) vs. M_{Post} = 31.29 (30.70), $t(45) = 2.48$, $p = .017$, $d = 0.24$), an AI chatbot (M_{Pre} = 33.51 (26.63) vs. M_{Post} = 26.75 (27.00), $t(53) = 3.85$, $p < .001$, $d = 0.25$), and a chatbot

acting as human (M_{Pre} = 38.44 (32.62) vs. M_{Post} = 31.41 (31.93), $t(86) = 4.20$, $p < .001$, $d = 0.22$), supporting hypothesis 1 (figure 2). The effect sizes were largest for the human, chatbot, and chatbot acting as human conditions. Results of the remaining metrics are reported in web appendix figure S5.

Additionally, we ran an exploratory test of whether state loneliness decreased after the interaction for participants in the chatbot acting as human condition who were successfully deceived, that is, who thought they were talking to another person, and not deceived, that is, who understood they were talking to an AI chatbot. For participants who were successfully deceived, we found that loneliness significantly decreased after the interaction (M_{Pre} = 42.70 (33.12) vs. M_{Post} = 32.83 (33.16), $t(31) = 2.85$, $p = .008$, $d = 0.30$). Similarly for the subset who were not deceived, we found that loneliness significantly decreased after the interaction (M_{Pre} = 35.96 (32.38) vs. M_{Post} = 30.58 (31.47), $t(54) = 3.15$, $p = .003$, $d = 0.17$).

Expectation Violation. We used a composite of the loneliness and social connection (reverse-coded) items to capture overall perceptions of whether the condition option would make people feel lonelier ($\alpha = 0.72$), as these measures are directly related—higher social connection generally reduces feelings of loneliness (Holt-Lunstad 2021). We compared this item before versus after the experience. In the web appendix, we replicated the analyses for loneliness and social connection separately and found similar results.

There was no significant expectation violation in loneliness for interacting with a human ($M_{Expected}$ = 36.32 (16.24) vs. M_{Actual} = 32.91 (25.62), $t(45) = 0.89$, $p = .376$, $d = 0.16$) or doing nothing ($M_{Expected}$ = 61.18 (21.46) vs. M_{Actual} = 63.71 (21.01), $t(57) = -1.01$, $p = .319$, $d = -0.12$). However, participants felt less lonely than they expected after watching a YouTube video ($M_{Expected}$ = 45.05 (18.61) vs. M_{Actual} = 36.81 (19.66), $t(36) = 5.09$, $p < .001$, $d = 0.43$), as well as after interacting with an AI chatbot ($M_{Expected}$ = 43.56 (17.11) vs. M_{Actual} = 34.46 (21.25), $t(53) = 4.13$, $p < .001$, $d = 0.47$). The expectation violation for the chatbot acting as human was marginally significant ($M_{Expected}$ = 37.11 (15.26) vs. M_{Actual} = 33.27 (21.27), $t(86) = 1.86$, $p = .066$, $d = 0.20$), supporting H5—figure 2. Again, we note that the effect sizes were largest for the AI chatbot and chatbot acting as human conditions. Also, although participants had low expectations for YouTube and chatbots, recall that only interacting with chatbots reduced stated loneliness. We also replicate the expectation violation result for AI chatbots, specifically, in another study (web appendix study S2) in which the effect was statistically significant.

In the web appendix, we report results for our other preregistered measures—entertainment, novelty, engagement, comfort, and interest—which compare participants' expectations before the interaction to their actual experience afterward. In the chatbot acting as human condition, participants reported higher-than-expected levels of entertainment, engagement, comfort, and interest. For comfort, we also found significant expectation violations in both the chatbot and human conditions. Finally, we also ran our preregistered moderation analysis (PROCESS Model 1; Hayes 2012) to see whether the expectation violation effect for loneliness, social connection, and comfort are moderated by attitudes toward AI, and we did not find significant moderation (web appendix).

In short, interacting with an AI companion improved their baseline loneliness levels on par only with interacting with another person, whereas a common technological alternative did not. Furthermore, participants underestimated the degree to which AI companions improved their loneliness relative to their true feelings after interacting with such AI (hypothesis 5). Future research can focus on reasons behind this misprediction, such as a general lack of familiarity with AI companions or more specific stereotypes about AI companions, like beliefs that AI companions are not capable of genuine understanding or providing emotional support.

STUDY 3

Study 3 aims to replicate the loneliness alleviating effects of an AI companion using a longitudinal design in which participants interact with the same chatbot daily for one week. We compared loneliness levels in participants before and after they interacted with an AI companion (experience condition) and contrasted these findings with a control group that did not engage with the AI (control condition), allowing us to directly assess the efficacy of AI companions in mitigating loneliness by comparing differences between the experience and control groups as well as changes within individuals over time. Given the findings of study 2, we hypothesized an immediate improvement in loneliness from the first day of interaction in the experience condition, although we did not set specific expectations for subsequent days. Additionally, we investigated whether consumers underestimate the efficacy of AI companions in reducing loneliness (prediction condition). We anticipated that consumers would likely underestimate the chatbot's capacity to lessen loneliness on the first day, although we did not have specific expectations for what predictions they would make for subsequent days.

Methods

This study was preregistered (https://aspredicted.org/BJD_JYZ). We recruited 1,088 participants from CloudResearch Connect and excluded 16 for failing a comprehension question, leaving 1,072 (M_{age} = 39.6, 47.3% female). Our goal was to enroll 200 participants in the

prediction condition, and 400 participants each in both the control and experience conditions, as these two conditions had a longitudinal design and we anticipated a 50% attrition rate. The prediction condition was not longitudinal and was only completed on day 1. Participants were randomly assigned to one of these conditions. Participants in the control and experience conditions were instructed to complete the study every day for seven days, and they were required to complete each day's survey by 12 AM local time. The survey for the following day became available at around 1 AM EST, and participants in earlier time zones needed to wait until it was the designated day in their time zone to proceed. Additionally, our chatbot app in the experience condition (explained further below) was designed to block users from entering the chatroom if they attempted to access it before the designated day in their time zone. If a participant failed to complete a session on its designated day, they were not invited to participate the following day. Attrition over seven days was much lower than we expected (likely because of the substantial bonus upon completion of the week-long study; see below). It amounted to 92 participants in the experience condition (23%) and 58 participants in the control condition (14%), leaving 922 participants in total (M_{age} = 40.1, 46.3% female). This difference in attrition (χ^2 (1, $N = 420 + 406$) = 10.29, p = .001) is likely due to the fact that the control condition required less time and effort than the experience condition, wherein participants had to interact with the chatbot. For example, participants in the control condition completed the study in 1.4 minutes on average, and participants in the experience condition completed it in 19.8 minutes (given the considerable cost of this study, we did not ask participants in the control condition to spend 15 minutes each day performing another task; see below). Below, we conduct a series of robustness checks, including a propensity score matching (PSM) analysis and one including dropouts also. These robustness checks confirm the conclusions of the preregistered analysis. Note also that differential attrition is irrelevant to the crucial pre–post loneliness difference in the experience condition. Overall, we had 246 participants for analysis in the prediction condition, 314 in the experience condition, and 362 in the control condition. Thirty-six percent had prior experience with chatbots. We ran this experiment between April 9, 2024, and April 15 2024.

Participants were paid $1 USD in the prediction and experience conditions, and $0.3 USD in the control condition, as the control condition took less time compared to the other two conditions. We stated to participants that they might be assigned to one of many conditions (varying in length and payment), and the time of each session ranges from 3 minutes over seven days for $0.3 in a single session to 20 minutes in a single session for $1. Participants were also notified that they would be awarded a $15 bonus after completing all seven days if they were assigned to one of the longitudinal conditions.

In the prediction condition, participants were asked to imagine interacting for 15 minutes with an AI companion every day for a week and were shown a screenshot of the AI companion app. On the next page, they were told, "In the next section, you will be asked to predict how lonely you would feel both before and after interacting with the chatbot, for each day of the 7 days." Next, participants were presented with the following text for each day of the study, using separate pages for each day: "Imagine it is Day X of interacting with the chatbot. Please answer the following questions about how lonely you would feel both before and after interacting with the AI companion for 15 minutes. For each statement, indicate the extent to which you agree that you would feel this way on Day X." Following this, participants reported their predictions of loneliness for both before and after imagining a 15-minute session with the AI companion, with each day's predictions entered on a new page.

In the control condition, participants were told, "As a reminder, in this longitudinal study, you will report your loneliness level every day for a week." Participants then answered the state loneliness questions once per day, which was the only task required in this condition. In the experience condition, participants were told, "As a reminder, in this longitudinal study, you will interact with a conversational AI companion every day for a week and will answer some questions before and after the interaction." Then, participants reported their state loneliness both before and after interacting for 15 minutes with the chatbot. For loneliness questions, participants in all conditions answered the same Three-Item UCLA Loneliness Scale as in study 2 (Hughes et al. 2004). In this condition, we used the same chatbot app as in the previous study except for the following changes: First, we utilized OpenAI's GPT-4 (gpt-4–0125-preview) because it was a more advanced model compared to GPT-3. Second, we implemented a memory feature that allows the chatbot to remember details from previous conversations with users. To do this, we periodically summarized the user's messages and integrated relevant information into future chats. By integrating this memory feature, the chatbot became capable of retaining user information for use in later conversations. Third, based on several pilots, we updated the model prompt to elicit chatbot responses that would be caring and friendly but not overly enthusiastic (see web appendix for the full prompt). Fourth, we implemented a check-in feature that prompts the chatbot to reach out to users if they have been inactive for 2 minutes. For this, we sent the following prompt to the chatbot: "The user did not send a message in the last 2 minutes. Check-in with the user, e.g., say 'Are you still

there?', or ask a question about the topic you were talking about." The chatbot then checked in with the user according to this prompt. Fifth, to prevent our OpenAI account from being banned due to the use of explicit language, we integrated OpenAI's moderation API (Markov et al. 2023). This API identifies and flags text containing explicit content. When such content is detected, we automatically replaced the flagged message with "[Harmful content]" before submitting it to GPT-4 to generate a response.

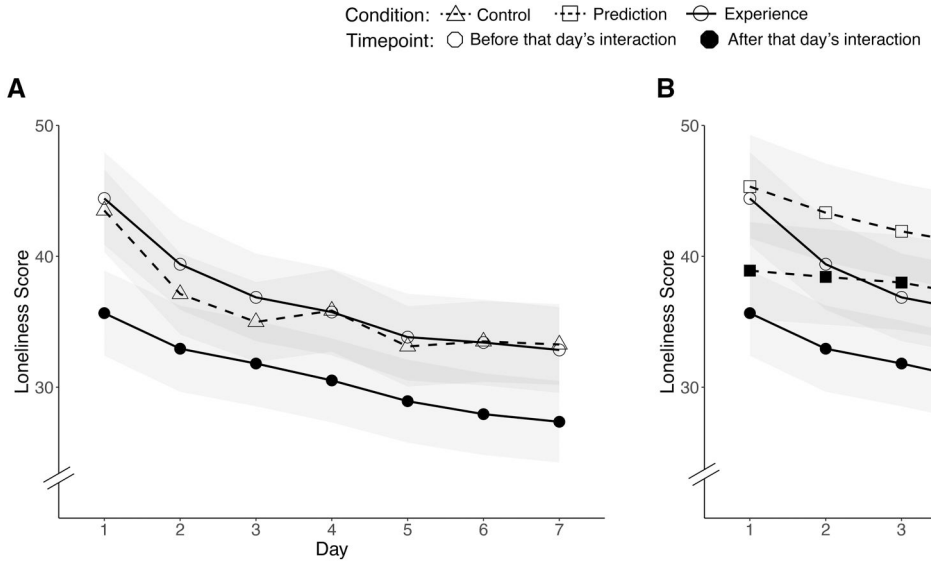
Participants in all conditions answered two comprehension checks on the first day: (1) "What was the topic of the questions you were asked? [Options: 'Loneliness,' 'Joint pain,' 'Nutritional advice']"; and (2) "On each day, what were you asked to predict/rate? [Options: 'How you (would) feel today/the same day,' 'How you (would) feel next month,' 'How you (would) feel next year']." We excluded 16 participants for failing either of these questions. Finally, participants indicated any prior experience with chatbots and completed the demographic questions on day 1. On the last day, participants in the experience condition answered the following questions about the chatbot: (1) "As you reflect on the last 7 days, how helpful was the chatbot for decreasing your loneliness?"; (2) "What aspects of the chatbot did not work well for you? Please provide specific examples or areas where you faced challenges"; and (3) "In what ways can we improve this chatbot to better support users like you? Feel free to suggest specific features, changes, or additions."

Results

We limited our analysis to participants who completed all seven days of the study in the longitudinal conditions because these were the participants who successfully fulfilled the study requirements. Following our preregistered analysis plan, we first ran a mixed-effects analysis of variance (ANOVA) on the experience condition, with loneliness as the dependent variable (DV), and timing (before vs. after interaction) and day (1–7) as the independent variables (IVs; i.e., we used the following model: Loneliness $\sim$ Timing $\times$ Day + (1 | Participant ID)). First, we found significant loneliness alleviation via the main effect of timing ($b = 7.61$, $p < .001$), as loneliness before interaction was significantly higher than loneliness after interaction when we aggregated the data over all days ($M_{\text{Before}} = 36.64$ vs. $M_{\text{After}} = 30.74$, $t(2197) = 20.15$, $p < .001$, $d = 0.20$). To further delineate daily changes in loneliness, we conducted paired t-tests comparing levels of loneliness before and after interaction with the chatbot for each individual day. We found that participants experienced a significant decrease in loneliness after each daily session with the chatbot supporting hypothesis 2 ($ps < .001$; table S10), and when comparing the postexperience loneliness with the control condition, loneliness levels were significantly lower on most days, partially supporting hypothesis 2 (figure 3A; more information in the next paragraph). We also found a main effect of day, indicating a gradual decrease in

FIGURE 3

RESULTS IN STUDY 3



NOTE.—Shadings indicate 95% confidence intervals. (A) The control and experience conditions are compared, and (B) the experience and prediction conditions are compared.

loneliness in the experience condition over the course of the week ($b = -0.92$, $p < .001$). Given that we also see this reduction in loneliness in the control condition ($b = -1.42$, $p < .001$; figure 3A), this is likely due to the longitudinal nature of the design. Supporting this, we also found that control versus before experience condition is not significant for all seven days ($ps > .337$; table S11). The gradual decrease in loneliness observed in both conditions might be attributed to participants perceiving the repetitive nature of the study, which involved daily check-ins, as possibly caring and supportive. Finally, we found a significant interaction between timing and day in the experience condition ($b = -0.43$, $p = .010$). However, this interaction effect was driven by day 1, as we did not see an interaction effect when we removed day 1 and reran the model ($b = -0.11$, $p = .566$); in other words, there was a particularly sharp drop in loneliness on the first day, with the subsequent six days showing similar-sized drops.

Second, in order to determine whether loneliness levels after experiencing the chatbot were lower than in the control condition, we ran the following ANOVA model on data from both the control condition and the "after" measurements from the experience condition: Loneliness $\sim$ Condition $\times$ Day + (1 | Participant ID). We found a main effect of both day ($b = -1.42$, $p < .001$) and condition ($b = -5.46$, $p = .015$) on loneliness, and there was no significant interaction ($b = 0.07$, $p = .455$). Specifically, loneliness was significantly lower after the chatbot interaction compared to the control condition on four out of seven days ($ps < .020$; table S12; figure 3A), marginally lower on day 2 ($M_{Control} = 37.13$ vs. $M_{After} = 32.94$, $t(662.2) = 1.82$, $p = .069$, $d = 0.14$) and day 5 ($M_{Control} = 33.12$ vs. $M_{After} = 28.94$, $t(666.7) = 1.86$, $p = .063$, $d = 0.14$), and directionally but not significantly lower on day 3 ($M_{Control} = 34.99$ vs. $M_{After} = 31.81$, $t(661.1) = 1.40$, $p = .163$, $d = 0.11$), partially supporting hypothesis 2. These small effect sizes (i.e., 0.11 and 0.14) suggest that the study may have been underpowered to detect significant differences on certain days. A post hoc power analysis indicates that detecting a significant effect of $d = 0.11$ with 80% power at $\alpha = 0.05$ would require approximately 1,298 participants per group. Even so, the consistent pattern of loneliness reduction in the intervention group, compared to the control group, suggests that the effects of the intervention are distinguishable from any general decrease in loneliness over time. Importantly, when we include only participants whose loneliness score was above the mean preinteraction score on the first day, loneliness was significantly lower after chatbot interactions than in the control condition on all days even without any bootstrapping ($ps < .01$, $ds > 0.28$), except that day 3 was marginally significant ($M_{Control} = 55.89$ (23.46) vs. $M_{After} = 51.17$ (25.02), $t(339.4) = 1.81$, $p = .071$, $d = 0.19$; table S13). This suggests that those with greater initial loneliness derive more consistent benefits from the intervention.

Third, in order to assess whether there was a difference in predicted versus actual drops in loneliness, we ran another ANOVA model, with the loneliness difference between before and after ratings on both prediction and experience conditions as the DV and condition and day as IVs; that is, we used the following model: Loneliness Difference $\sim$ Condition $\times$ Day + (1 | Participant ID). We found a main effect of day ($b = -0.43$, $p < .001$), indicating that the before and after loneliness difference generally decreased over the days. However, this interaction effect was again largely driven by day 1, as we did not see a main effect of day when we removed day 1 and reran the model ($b = -0.11$, $p = .295$); in other words, the difference between before versus after ratings on the prediction and experience conditions was the largest on the first day, with the subsequent six days showing similar-sized ratings. The main effect of condition was not significant ($b = -1.72$, $p = .174$), and there was no significant interaction effect ($b = 0.05$, $p = .731$). Additionally, for each day, there was no significant difference in loneliness between the prediction and experience conditions ($ps > .146$), although the loneliness reduction was consistently numerically higher in the experience condition. One possibility is that, compared to study 2, participants in the current study might have had higher expectations regarding the capabilities of chatbots due to the increased popularity of ChatGPT. Further, when we aggregated the data over all seven days, we found that participants significantly underestimated the chatbot's ability to reduce loneliness ($M_{Prediction} = 4.37$ vs. $M_{Experience} = 5.91$, $t(3177.7) = -2.96$, $p = .003$, $d = -0.10$; figure 3B).

Overall, participants reported a decrease in loneliness after their interactions with the chatbot. This is also reflected in comments provided through CloudResearch Connect's comment feature: "this was a very interesting survey and i think it would help people who are really lonely and need someone to talk to"; "I am really enjoying my talks with Jessie. Its so easy and it feels really amazing to have someone (or something. . .I guess?) listen . . . and the responses I get are perfect, to be honest"; "It's funny. I wasn't sure how I was going to feel about this, talking every day to the AI about whatever comes to mind for 15 minutes, but now it's become a rather pleasant routine. I could see where this would really benefit people who were feeling isolated. . . ."

Next, as an exploratory analysis, we investigated whether greater engagement, measured as number of turns and mean number of words sent by participants on each conversational turn, was associated with a greater decrease in loneliness. For this, we ran the following ANOVA model on the experience condition: Loneliness Difference $\sim$ No. Messages + No. Words + (1 | Participant ID). We found a positive main effect of the number of messages on the loneliness difference ($b = 0.08$, $p = .046$); that is, participants who sent more messages to the chatbot experienced a higher decrease in loneliness. Although the exact

reason for this effect is unclear, one possibility is that more interactions enable the chatbot, possibly aided by its memory feature, to tailor its responses more effectively to the user. To examine whether the reduction in loneliness was consistent among participants who sent fewer messages, we conducted a supplementary analysis excluding 179 participants who sent fewer messages than the overall mean (17.79), leaving 270 participants. Even within this lower-engagement subset, loneliness significantly decreased from pre- to postinteraction (all $ps < .001$; table S14). We did not find a main effect of the number of words on the loneliness difference ($b = 0.03$, $p = .331$).

Finally, we conducted two additional analyses to assess the robustness of the results. Due to the higher attrition rate in the experience condition (23%, 92 out of 406) compared to the control condition (14%, 59 out of 421), we employed PSM using the nearest neighbors method to address possible selection bias (Austin 2011). In this method, participants from the experience condition were matched with participants from the control condition based on closely similar demographics (web appendix). This method allowed selecting participants who were demographically similar between control and experience conditions. After applying PSM, our findings were consistent with those obtained from the original sample. To further corroborate this conclusion, in the web appendix, we report another replication of the analyses including participants who did not complete all seven days of the study. These results were also similar, confirming the robustness of our findings.

STUDY 4

Study 4 investigates what types of features of the chatbot reduce loneliness and whether AI companions reduce loneliness more than generalist AI assistants and highly constrained chatbots. In the market, people can interact with various types of chatbots. For example, ChatGPT lacks anthropomorphic cues and uses a system prompt like "You are a helpful assistant," whereas Replika anthropomorphizes the chatbot and emphasizes empathy. In this study, we compare these two chatbot types to understand their differential impacts on loneliness. Participants interacted with three different chatbots: (1) the same AI companion as in study 3, (2) a generalist AI assistant that assists participants on various topics without offering emotional responses, and (3) a limited AI assistant that was only able to help with unit conversion, basic arithmetic, and grammar—this was the control condition (all chatbots were based on the same LLM: OpenAI's GPT-4). We hypothesized a decrease in loneliness in the AI companion condition but were agnostic about the AI assistant condition. We also predicted that the loneliness difference would be higher in the AI companion condition compared to the other conditions and that this effect would be primarily driven by the feeling of being heard by the chatbot, although we also investigated whether performance perceptions play a mechanistic role.

Methods

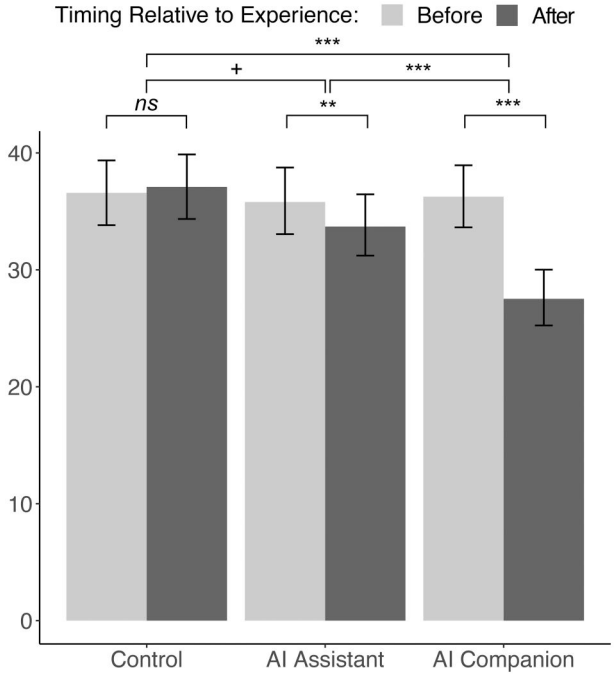
This study was preregistered (https://aspredicted.org/XCY_LLD). We recruited 1,479 participants from CloudResearch's Connect and excluded 98 for failing a comprehension question, leaving 1,381 ($M_{age} = 39.9$, 57.1% female). We aimed to hire 500 participants in all conditions, and participants were randomly assigned to one of three conditions: AI companion, AI assistant, and control (i.e., the limited functionality chatbot). Prior experience with chatbots was reported by 52.9% of participants. We ran this experiment on May 16, 2024. All participants were paid $2.75 USD.

All participants were asked to complete the same loneliness questions as in the previous study before interacting with the chatbot for 15 minutes. After interacting with the chatbot, participants completed the same loneliness scale in addition to ratings of feeling heard ($\alpha = 0.96$), measured with three items (Roos et al. 2023; Zielinski and Veilleux 2018) such as "The chatbot put itself in my shoes," and chatbot performance ($\alpha = 0.84$) measured with five items (Borsci et al. 2022) such as "The chatbot was able to keep track of context." All items were measured with 100-point scales, ranging from "strongly disagree" to "strongly agree." For all questions, see table S16 in the web appendix. Next, participants completed the following comprehension checks: (1) "What was the topic of the questions you were asked? [Options: 'Loneliness', 'Joint pain', 'Nutritional advice']," and (2) "What were you asked to rate? [Options: 'How you feel today/next month/next year']." Finally, participants answered a question about AI capability, indicated any prior experience with chatbots, and completed the demographic questions.

Chatbots in all conditions were the same, except for the prompts with which they were seeded. Additionally, the name of the chatbot differed across conditions: It was "AI assistant" in the AI assistant and control conditions and "Jessie" as before in the AI companion condition. In AI assistant and control conditions, the chatbot's writing notification was also shown as "Processing your request, please wait," instead of "Jessie is writing." Each message bubble also contained the text "Message generated by AI system" at the lower left. Participants in the AI companion condition interacted with the same chatbot as in the previous study (study 3). Those in the AI assistant condition interacted with a generalist chatbot that was able to assist participants with various topics without offering emotional responses. The chatbot in this condition was prompted to provide assistance without personal interaction, maintain formal and precise language, and deliver concise, task-focused responses (see web appendix for the full prompt).

FIGURE 4

RESULTS IN STUDY 4



NOTE.—Bars indicate mean loneliness. ***$p < .001$, **$p < .01$, +$p < .1$. AI, artificial intelligence.

Participants in the control condition interacted with a rudimentary chatbot, which was only able to assist with basic grammar, unit conversion, and basic arithmetic. This chatbot was prompted to perform these limited tasks, decline unrelated requests, and maintain concise, emotionless responses (see web appendix for the full prompt).

Results

Following our preregistered analysis plan, we first ran paired *t*-tests comparing loneliness before versus after the experience and found that loneliness was significantly lower after the experience in both AI companion (M_{Before} = 36.26 (SD = 30.36); M_{After} = 27.53 (26.87); $t(491)$ = 10.61, $p < .001$, $d = 0.30$) and AI assistant (M_{Before} = 35.80 (29.86); M_{After} = 33.70 (29.2); $t(440)$ = 2.62, p = .009, $d = 0.07$) conditions but not in the control condition (M_{Before} = 36.58 (29.75); M_{After} = 37.09 (30.87); $t(447)$ = −0.57, p = .571, $d = -0.02$; figure 4).

Second, we conducted a one-way ANOVA to test for differences in loneliness reduction across the three conditions. We found a significant effect of condition on loneliness reduction ($F(2, 1378) = 32.99$, $p < .001$, $\eta^2 = 0.05$; figure 4). Next, we conducted post hoc pairwise comparisons using Tukey's honestly significant difference (HSD) test. Loneliness reduction in the AI companion condition was significantly greater compared to both the AI assistant condition ($M_{\text{AI Companion}}$ = 8.73 (18.25); $M_{\text{AI Assistant}}$ = 2.10 (16.78); $p < .001$; 95% confidence interval [CI; 3.86–9.41]) and the control condition ($M_{\text{AI Companion}}$ = 8.73 (18.25); M_{Control} = −0.51 (19.04); $p < .001$; 95% CI [6.47–12.01]). The difference in loneliness reduction between the AI assistant condition and the control condition was only marginally significant ($M_{\text{AI Assistant}}$ = 2.10 (16.78); M_{Control} = −0.51 (19.04); p = .080; 95% CI [0.24–5.45]).

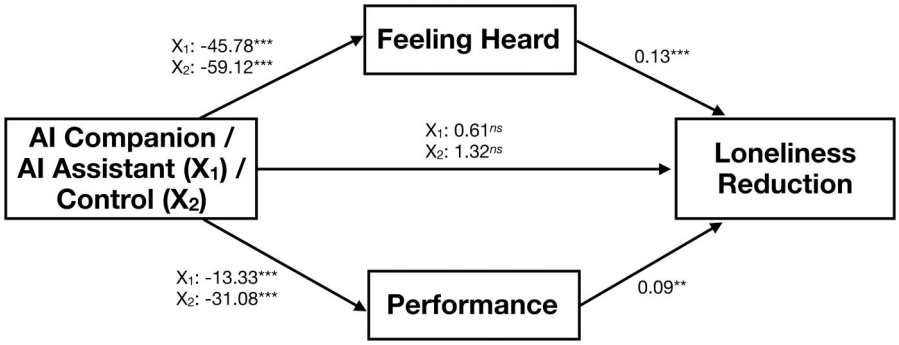
Third, we conducted an ANOVA to test for differences in feeling heard across conditions. We found a significant effect ($F(2, 1378) = 899.2$, $p < .001$, $\eta^2 = 0.57$). Post hoc analyses found that feeling heard in the AI companion condition was significantly greater compared to both the AI assistant condition ($M_{\text{AI Companion}}$ = 70.63 (22.83); $M_{\text{AI Assistant}}$ = 24.85 (24.59); $p < .001$; 95% CI [42.31–49.24]) and the control condition ($M_{\text{AI Companion}}$ = 70.63 (22.83); M_{Control} = 11.51 (19.93); $p < .001$; 95% CI [55.67–62.57]). Feeling heard was also significantly higher in the AI assistant condition compared to control ($M_{\text{AI Assistant}}$ = 24.85 (24.59); M_{Control} = 11.51 (19.93); $p < .001$; 95% CI [9.80–16.89]).

Fourth, we conducted an ANOVA to test for differences in performance across all three conditions and found a significant effect of condition on performance ($F(2, 1378)$ = 315, $p < .001$; $\eta^2 = 0.31$). Post hoc analyses found that performance in the AI companion condition was significantly greater compared to both the AI assistant condition ($M_{\text{AI Companion}}$ = 82.25 (15.55); $M_{\text{AI Assistant}}$ = 68.92 (19.18); $p < .001$; 95% CI [10.41–16.25]) and the control condition ($M_{\text{AI Companion}}$ = 82.25 (15.55); M_{Control} = 51.17 (21.99); $p < .001$; 95% CI [28.17–33.99]). We also found that performance was significantly higher in the AI assistant condition compared to control ($M_{\text{AI Assistant}}$ = 68.92 (19.18); M_{Control} = 51.17 (21.99); $p < .001$; 95% CI [14.76–20.73]).

Fifth, we ran a mediation model (PROCESS Model 4; Hayes 2012) with AI companion/AI assistant/control as the multicategorical IV, feeling heard and performance as mediators, and loneliness reduction as the DV (figure 5). We set the AI companion condition as the reference group and compared it to the AI assistant condition (X_1) and control conditions (X_2; Montoya and Hayes 2017). We found that feeling heard mediated the effect of loneliness reduction both relative to the AI assistant ($b = -6.08$, standard error [SE] = 1.22, 95% CI [−8.51 to −3.72]) and control conditions ($b = -7.86$, SE = 1.57, 95% CI [−10.97 to −4.82]), indicating that the effect of loneliness reduction was driven by feeling heard (figure 5), supporting hypothesis 3a. As for performance, we found that performance mediated the effect of loneliness reduction relative to both the control condition ($b = -1.16$, SE = 0.45, 95% CI

FIGURE 5

MEDIATION DIAGRAM IN STUDY 4



NOTE.—***$p < .001$, **$p < .01$. AI, artificial intelligence.

[−2.08 to −0.30]) and the AI assistant condition ($b = -2.70$, SE = 1.03, 95% CI [−4.76 to −0.73]), indicating that the reduction in loneliness for the AI companion versus AI assistant and control conditions was influenced by both feeling heard and performance. Notably, when comparing feeling heard to performance in the control condition, the coefficient for feeling heard ($b = -7.86$) was more than six times larger than that for performance ($b = -1.16$). Similarly, in the AI assistant condition, the coefficient for feeling heard was more than twice as large ($b = -6.08$ vs. −2.70), suggesting that feeling heard played a larger role than performance in both the control and AI assistant conditions. Because multicategorical mediation analysis does not allow statistically comparing mediators, we ran two additional mediation models in which the AI companion condition was compared to the AI assistant and control conditions, respectively. Feeling heard was a significantly stronger mediator than performance in both comparisons: AI assistant ($b = -3.20$, SE = 1.49, 95% CI [−6.18 to −0.28]) and control ($b = -5.05$, SE = 1.54, 95% CI [−8.10 to −2.09]), supporting hypothesis 3b.

Finally, we ran an exploratory serial mediation model (PROCESS Model 6; Hayes 2012), wherein "feeling heard" influences "performance," which in turn affects "loneliness reduction" (i.e., condition → feeling heard → performance → loneliness reduction). We found a significant indirect effect, both for the AI assistant ($b = -1.69$, SE = 0.65, 95% CI [−2.97 to −0.45]) and for control conditions ($b = -2.18$, SE = 0.83, 95% CI [−3.82 to −0.58]). However, the indirect effect of the parallel model with feeling heard as the mediator had an effect size that was several times larger compared to this serial model (3.5 times higher for AI assistant and 3.6 times higher for control). This simpler model also fit significantly better compared to the serial model for the control condition ($b = -5.09$, SE = 1.50, 95% CI [−8.07 to −2.20]), and it was not significantly different from the model for the AI assistant condition ($b = -2.50$, SE = 1.64, 95% CI [−5.75 to 0.71]). Thus, the results are consistent with "feeling heard" and "performance" independently reducing loneliness. Intuitively, it is feeling heard, rather than performance, that directly predicts loneliness reduction. We also examined discriminant validity between the two constructs using the heterotrait–monotrait (HTMT) ratio. The HTMT ratio compares the average correlations between items of different constructs to the average correlations between items within the same construct, providing an estimate of discriminant validity—that is, whether two constructs are empirically distinct. The HTMT ratio between feeling heard and performance was 0.69, well below the commonly accepted thresholds of 0.85 or 0.90 (Henseler, Ringle, and Sarstedt 2015), indicating that the constructs are related but sufficiently distinct.

STUDY 5

Study 5 tests whether AI companions reduce loneliness more effectively than journaling. In pilot studies 1 and 2, we found that journaling involves *higher* self-disclosure and similar levels of distraction compared to AI companions and that watching YouTube videos involves *higher* distraction (see web appendix for all details). Specifically, in pilot study 1, participants either interacted with an AI companion or engaged in a journaling task and then answered self-disclosure questions (Cayanus and Martin 2004), such as "I shared personal thoughts, feelings, or experiences during this interaction." We found that journaling produced significantly higher self-disclosure than interacting with AI companions. In pilot study 2, participants either watched YouTube videos, journaled, or interacted

with the AI companion; they then responded to distraction questions (Lopez et al. 2023), such as "During [the activity], I often found myself distracted by other thoughts." We found that journaling and AI companions did not differ on distraction, whereas watching YouTube produced significantly higher distraction.

Because AI companions reduced loneliness more than watching YouTube videos in study 2, this already speaks against distraction as an alternative mechanism. Study 5 further examines self-disclosure and distraction as alternative mechanisms (hypothesis 4) by comparing loneliness alleviation between AI companions and journaling. We also investigate whether the loneliness-alleviating benefits of AI companions persist when anthropomorphic cues are minimized (see web appendix for details). Finally, study 5 tests the robustness of the loneliness-alleviating benefits of AI companions. The pre–post design that we used in studies 1–4 for the measurement of loneliness is the standard approach to test the effect of interventions in the loneliness literature because it affords more precise estimates of the effect of interventions (by calibrating the effect at the individual level) and because it enables assessing successful randomization. However, this approach presents the risk of reducing the external validity of the test by raising the salience of the loneliness construct before interacting with the chatbot in a way that would not normally occur in naturalistic situations. It is therefore possible that this feature of the design contributed to the loneliness reduction documented in previous studies. To address this concern, we measured loneliness only after the interaction and embedded loneliness questions among decoy questions to obscure the study's purpose from participants.

Methods

This study was preregistered (https://aspredicted.org/pfs9-t5q7.pdf). Based on the power analysis from a pilot study, we recruited 2,228 participants from CloudResearch's Connect (M_{age} = 40.7, 57.6% female). We aimed to hire 750 participants in each condition, and participants were randomly assigned to one of three conditions: AI companion, journaling, and control (i.e., the limited functionality chatbot from study 4). Prior experience with chatbots was reported by 64.9% of participants. We ran this experiment on December 18, 2024. All participants were paid $3 USD. After the intervention, participants answered loneliness items embedded among eight decoy questions (e.g., about lighting, temperature, or confusion). At the end of the survey, we asked them to type what they thought the study was testing. See web appendix for more details.

HTMT ratios below 0.46 indicated that the loneliness measures were distinct from the decoy questions (Henseler et al. 2015). Only 2% of participants suspected the study was about loneliness, indicating that our decoy items successfully obscured the study's purpose. We replicated the analyses in the web appendix after excluding these participants and found similar results.

Results

A one-way ANOVA revealed a significant effect of condition on loneliness ($F(2, 2225) = 36.92$, $p < .001$). Post hoc tests (Tukey's HSD) indicated that loneliness was significantly lower in the AI companion condition than both the control condition ($M_{\text{AI Companion}} = 21.84$ (24.70); $M_{\text{Control}} = 30.71$ (27.67); $p < .001$; 95% CI [5.44–12.29]) and the journaling condition ($M_{\text{Journaling}} = 33.65$ (30.92); $p < .001$; 95% CI [8.46–15.16]). There was no significant difference between the control condition and the journaling condition ($p = .109$; 95% CI [−0.48 to 6.36]). These results suggest that the benefit of AI companions extends beyond self-disclosure and distraction (hypothesis 4). Web appendix study S3 replicates this finding of lower loneliness for AI companions in a design comparing AI companion versus control and no decoy questions.

GENERAL DISCUSSION

We considered whether apps that utilize AI to provide consumers with synthetic companions can reduce loneliness. Study 1 found correlational evidence that AI companions alleviate loneliness. Study 2 found that AI companions successfully alleviate loneliness on par only with interacting with another person, whereas watching YouTube videos or doing nothing do not. Furthermore, participants underestimate the degree to which AI companions improve their loneliness relative to their true feelings after interacting with such AI. Study 3 found that AI companions consistently provided momentary reductions in loneliness after use over the course of a week. The most significant reduction occurred on the first day, followed by stable decreases in loneliness on subsequent days. This suggests that the initial interaction with the AI companion has a pronounced impact, which quickly stabilizes over time as participants acclimate to their AI companions. However, these effects did not accumulate or persist in the absence of interaction. That is, the intervention reliably provided short-term relief but did not lead to a sustained lowering of loneliness across the week. Study 4 provided evidence that feeling heard and performance are significant mediators of the loneliness-alleviating effects of AI companions, with feeling heard exerting the greater influence in reducing loneliness. Study 5 demonstrated that AI companions reduce loneliness more than journaling—an activity involving higher self-disclosure (pilot study 1) and lower distraction (pilot study 2), even when participants do not know the study purpose because loneliness is measured amid decoy items.

Theoretical Contributions

We make several contributions. First, a number of reviews on loneliness and mental health have noted the need for evidence on new technological solutions (Shrum et al. 2023; Veronese et al. 2021). In this space, most work is correlational and qualitative (Maples et al. 2024; Ta et al. 2020), and the one exception focused on an extreme sample of older patients with serious mental illness during the pandemic, using a highly rule-based chatbot (Chou et al. 2024). We provide rigorous experimental evidence using state-of-the-art LLM-based AI companions to causally isolate the effectiveness of AI companions in reducing loneliness compared to other technological solutions and control conditions.

Although classic theories of social support (Baumeister and Leary 2017) emphasize the role of human relationships in satisfying the need for connection, our findings suggest that even brief, synthetic interactions can provide emotional relief. Whereas it is true that people have traditionally satisfied their need to belong by forming and sustaining social bonds with other humans, our results suggest that AI companions—despite lacking various traits of human beings—can provide benefits characteristics of social bonds as well. However, it is important to clarify that the effects observed in our studies are momentary in nature. That is, AI companions consistently alleviated loneliness immediately after each interaction, but we did not observe evidence that these reductions persisted over time before the interaction with the AI companion. These findings highlight the distinction between temporary emotional relief and lasting changes in baseline loneliness, suggesting that AI companions may offer acute support rather than enduring transformation.

Second, we contribute to understanding which features of chatbots lead to alleviation of loneliness (Merrill Jr. et al. 2022) by leveraging insights from psychological studies (Itzchakov et al. 2023; Kahlon et al. 2021; Myers 2000; Reis et al. 2017) and a growing literature in human–computer interaction on the role of "feeling heard" on relationships with chatbots (Boucher et al. 2021; De Gennaro et al. 2020; Leite et al. 2013). Although prior research has shown that consumers feel heard when companies engage with them in brand communities—enhancing brand trust (Bang et al. 2018)—there has been no work examining the effect of feeling heard by products themselves, such as AI companions. We find that using prompting to ensure that the AI is friendly and caring improves the sense that users feel heard, relative to general assistants without these capabilities, and that feeling heard explains levels of loneliness reduction.

Third, we contribute to previous findings showing people underestimate how much they enjoy socially interacting with other human strangers (Epley and Schroeder 2014; Kardas et al. 2022). We find that people also make a similar forecasting error about their interactions with AI.

Practical Implications

Many AI companions available on the market advertise loneliness alleviation as a value proposition, but to the best of our knowledge, our studies are the first to rigorously and causally assess whether this is the case. This finding is not only relevant to AI companion apps but also mental health apps that are increasingly incorporating "talk therapy" as part of their offerings. As shown in study 4, even AI assistants may alleviate loneliness to an extent. Finally, the results document the benefits of building "generalist" LLM-based chatbots with empathic features designed to make consumers feel heard. For example, Inflection AI's Pi was a chatbot explicitly designed to be, and marketed as being, a friendly conversationalist. From a societal point of view, the promising results found here suggest that AI companions could be a scalable tool for loneliness reduction against the backdrop of a severe loneliness crisis. Whether chatbots can help reduce loneliness has recently been the object of intense debate (Marriott and Pitardi 2024), and rigorous empirical evidence in this area was sorely needed. In a series of tightly controlled and high-powered experimental studies, we find compelling evidence that AI companions can indeed reduce momentary feelings of loneliness, at least at the time scales of a day and a week. These descriptive results pave the way for future consumer research on this topic.

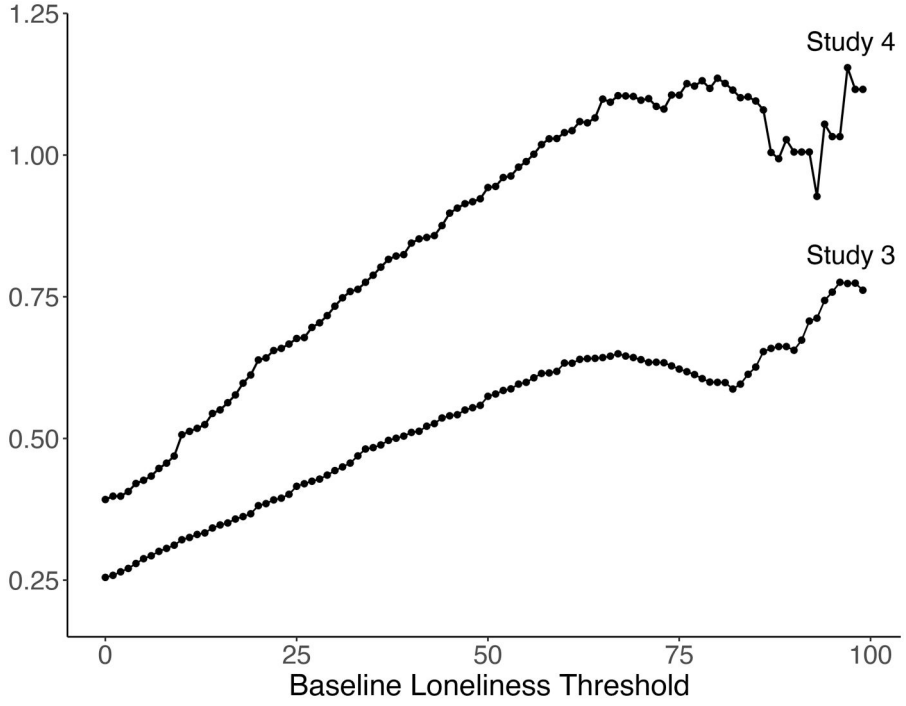
Although our findings demonstrate that AI companions can provide reductions in loneliness, we caution against interpreting these effects as evidence of long-term psychological improvement. Our studies examined whether repeated use of AI companions over the course of a week leads to lasting improvements in loneliness and found no evidence that reductions persist beyond the immediate interaction. Although our research is motivated by the public health significance of loneliness, it is not intended to inform clinical recommendations. We do not suggest that AI companions can or should replace human relationships or professional mental health care. Future research is needed to explore how AI companions might be integrated responsibly into broader ecosystems of support without undermining users' social development.

Potential Moderating Factors

Is the loneliness reduction effect moderated by participants initial loneliness levels? We explored this question across studies 2–4, finding that those with higher baseline loneliness experienced significantly greater reductions in loneliness after interacting with the AI companion in all studies (figure 6; for linear regression results, see "Loneliness Reduction and Baseline Loneliness Levels" sections in the web appendix). This pattern aligns with the social buffer hypothesis (Cohen and Wills 1985), which posits that individuals with limited social support derive greater benefits from supportive interventions.

FIGURE 6

EFFECT SIZE VERSUS BASELINE LONELINESS THRESHOLD ACROSS STUDIES 3 AND 4



NOTE.—The x-axis depicts different baseline loneliness thresholds; for each threshold, we included participants whose initial (i.e., before interaction) loneliness score was above that value. The y-axis indicates the effect size for the comparison of loneliness before versus after interaction. Thresholds were included only if both groups (pre- and postinteraction) contained at least 10 participants to ensure reliable effect size calculations. In study 4, we only included participants in the AI companion condition. Although the direction of the effect in study 2 was consistent with these findings, we do not report that analysis in figure 6 due to smaller cell sizes in that study.

Although our findings suggest that AI companions alleviate loneliness, the heightened impact on lonelier individuals could also raise concerns about further isolating them from human interaction. There is a potential risk that long-term reliance on AI companions may create unwanted dependencies, which could be harmful for users (Valenzuela et al. 2024). Such dependencies might lead individuals to substitute human relationships with AI interactions, exacerbating social isolation. Whether these apps have long-term effects on loneliness and mental health is an open question.

Another potential moderator is views about chatbots as well as anticipated social stigma from using them, as suggested by our misprediction results. Like other mental health issues (Henderson et al. 2014), loneliness is difficult to admit publicly, and alleviating loneliness with the use of AI might compound the shame one feels in admitting it. These anticipated social exclusion aftereffects of using an AI companion could potentially limit the efficacy of AI companions by reducing willingness to engage fully with the app, moderating its potential to alleviate loneliness. It could also moderate loneliness reduction by limiting utilization of AI companions in the first place, suggesting that loneliness alleviation is unrealized unless these adoption barriers are overcome.

Beyond social stigma, recent research also suggests that people's beliefs about chatbots may play a role. One study finds that people generally prefer to not use AI companions for relationship purposes because they view it as incapable of *mutual caring* (De Freitas et al. 2024a). The antecedents of this belief are the views that AI companions are incapable of understanding and feeling empathy. It is possible that similar stereotypes drive the misprediction results observed in the current studies.

Limitations and Future Research

Beyond questions about moderating factors, future research should also explore further which features of chatbots lead consumers to feel heard and, more broadly, what other psychological processes might contribute to loneliness alleviation. Ultimately, feeling heard might reduce loneliness because people have an innate "need to belong" in a group or community given that such belonging increases our chances of survival and success as a species (Baumeister

and Leary 2017). Another area for future research is to explore the specific conversational elements that drive loneliness reduction in interactions with AI companions. Although study 3 found that a greater number of conversational turns is associated with greater reductions in loneliness, it remains unclear which features within these turns contribute most to these effects. Finally, our studies were conducted within a specific cultural context (the United States), which may influence perceptions and effectiveness of AI companions. The generalizability of these findings across different cultures warrants further investigation.

DATA COLLECTION STATEMENT

The data for study 1 were collected from publicly available user reviews on the App Store. The third author scraped the data on January 24, 2024, for all apps except ChatGPT, which was scraped on February 4, 2024, and Wysa, which was scraped on November 27, 2024. The third author collected the data for study 2 from MTurk between April 5, 2023, and April 7, 2023. The third author collected the data for study 3 from CloudResearch Connect between April 9, 2024, and April 15, 2024. The third author collected the data for study 4 from CloudResearch Connect on May 16, 2024. The third author collected the data for study 5 from CloudResearch Connect on December 18, 2024. The first three authors jointly analyzed the data in all studies. All data are currently stored in a project directory on the Open Science Framework: https://osf.io/hf9xe/.

REFERENCES

American Psychological Association (2012), "Understanding Psychotherapy and How It Works." Last Accessed February 1, 2024. https://www.apa.org/topics/psychotherapy/understanding.

Araujo, Theo (2018), "Living up to the Chatbot Hype: The Influence of Anthropomorphic Design Cues and Communicative Agency Framing on Conversational Agent and Company Perceptions," *Computers in Human Behavior*, 85 (8), 183–9.

Austin, Peter C. (2011), "An Introduction to Propensity Score Methods for Reducing the Effects of Confounding in Observational Studies," *Multivariate Behavioral Research*, 46 (3), 399–424.

Bang, Juyoung, Seounmi Youn, James Rowean, Michael Jennings, and Manila Austin (2018), "Motivations for and Outcomes of Participating in Research Online Communities," *International Journal of Market Research*, 60 (3), 238–56.

Banks, Marian R., Lisa M. Willoughby, and William A. Banks (2008), "Animal-Assisted Therapy and Loneliness in Nursing Homes: Use of Robotic Versus Living Dogs," *Journal of the American Medical Directors Association*, 9 (3), 173–7.

Barbieri, Francesco, Luis Espinosa Anke, and Jose Camacho-Collados (2021), "Xlm-T: Multilingual Language Models in Twitter for Sentiment Analysis and Beyond," in *Proceedings of the Thirteenth Language Resources and Evaluation Conference*, Marseille, France, June 20–5, European Language Resources Association, 258–66.

Baumeister, Roy F. and Mark R. Leary (2017), "The Need to Belong: Desire for Interpersonal Attachments as a Fundamental Human Motivation," in *Interpersonal Development*, ed. Roy F. Baumeister and Mark R. Leary, London, UK: Routledge, 57–89.

Beaver, Kelly (2021), "Loneliness on the Increase Worldwide, But an Increase in Local Community Support." Last Accessed February 1, 2024. https://www.ipsos.com/en-cn/loneliness-increase-worldwide-increase-local-community-support.

Bergner, Anouk S., Christian Hildebrand, and Gerald Häubl (2023), "Machine Talk: How Verbal Embodiment in Conversational AI Shapes Consumer–Brand Relationships," *Journal of Consumer Research*, 50 (4), 742–64.

Bloomberg (2023), "Generative AI to Become a $1.3 Trillion Market by 2032, Research Finds." Last Accessed February 1, 2025. https://www.bloomberg.com/company/press/generative-ai-to-become-a-1-3-trillion-market-by-2032-research-finds/

Borsci, Simone, Alessio Malizia, Martin Schmettow, et al. (2022), "The Chatbot Usability Scale: The Design and Pilot of a Usability Scale for Interaction with AI-Based Conversational Agents," *Personal and Ubiquitous Computing*, 26 (1), 95–119.

Boucher, Eliane M., Nicole R. Harake, Haley E. Ward, et al. (2021), "Artificially Intelligent Chatbots in Digital Mental Health Interventions: A Review," *Expert Review of Medical Devices*, 18 (Supp1), 37–49.

Brakus, Joško J., Bernd H. Schmitt, and Lia Zarantonello (2009), "Brand Experience: What Is It? How Is It Measured? Does It Affect Loyalty?," *Journal of Marketing*, 73 (3), 52–68.

Brown, Tom, Benjamin Mann, Nick Ryder, et al. (2020), "Language Models Are Few-Shot Learners," *Advances in Neural Information Processing Systems*, 33, 1877–901.

Cacioppo, John T. and Stephanie Cacioppo (2018), "Loneliness in the Modern Age: An Evolutionary Theory of Loneliness (Etl)," in *Advances in Experimental Social Psychology*, Vol. 58, ed. J. M. Olson, Cambridge, MA: Elsevier, 127–97.

Castelo, Noah, Johannes Boegershausen, Christian Hildebrand, and Alexander P. Henkel (2023), "Understanding and Improving Consumer Reactions to Service Bots," *Journal of Consumer Research*, 50 (4), 848–63.

Castelo, Noah, Maarten W. Bos, and Donald R. Lehmann (2019), "Task-Dependent Algorithm Aversion," *Journal of Marketing Research*, 56 (5), 809–25.

Cayanus, Jacob L. and Matthew M. Martin (2004), "An Instructor Self-Disclosure Scale," *Communication Research Reports*, 21 (3), 252–63.

Chaves, Ana Paula and Marco Aurelio Gerosa (2021), "How Should My Chatbot Interact? A Survey on Social Characteristics in Human–Chatbot Interaction Design," *International Journal of Human–Computer Interaction*, 37 (8), 729–58.

Chen, Shu-Chuan, Wendy Moyle, Cindy Jones, and Helen Petsky (2020), "A Social Robot Intervention on Depression, Loneliness, and Quality of Life for Taiwanese Older Adults in Long-Term Care," *International Psychogeriatrics*, 32 (8), 981–91.

Chou, Ya-Hsin, Chemin Lin, Shwu-Hua Lee, Yen-Fen Lee, and Li-Chen Cheng (2024), "User-Friendly Chatbot to Mitigate the Psychological Stress of Older Adults During the Covid-19 Pandemic: Development and Usability Study," *JMIR Formative Research*, 8, e49462.

Chung, Minjee, Eunju Ko, Heerim Joung, and Sang Jin Kim (2020), "Chatbot E-Service and Customer Satisfaction Regarding Luxury Brands," *Journal of Business Research*, 117, 587–95.

Cohen, Sheldon and Thomas A. Wills (1985), "Stress, Social Support, and the Buffering Hypothesis," *Psychological Bulletin*, 98 (2), 310–57.

Croes, Emmelyn A. J. and Marjolijn L. Antheunis (2021), "Can We Be Friends with Mitsuku? A Longitudinal Study on the Process of Relationship Formation between Humans and a Social Chatbot," *Journal of Social and Personal Relationships*, 38 (1), 279–300.

Crolic, Cammy, Felipe Thomaz, Rhonda Hadi, and Andrew T. Stephen (2022), "Blame the Bot: Anthropomorphism and Anger in Customer–Chatbot Interactions," *Journal of Marketing*, 86 (1), 132–48.

Darcy, Alison, Jade Daniels, David Salinger, Paul Wicks, and Athena Robinson (2021), "Evidence of Human-Level Bonds Established with a Digital Conversational Agent: Cross-Sectional, Retrospective Observational Study," *JMIR Formative Research*, 5 (5), e27868.

De Freitas, Julian, Stuti Agarwal, Bernd Schmitt, and Nick Haslam (2023), "Psychological Factors Underlying Attitudes toward AI Tools," *Nature Human Behaviour*, 7 (11), 1845–54.

De Freitas, Julian and Nicole Tempest Keller (2022), *Replika AI: Monetizing a Chatbot*, Boston, MA: Harvard Business Publishing.

De Freitas, Julian, Zeliha Oğuz-Uğuralp, Ahmet Kaan Uğuralp, and Stefano Puntoni (2024a), "Why Most Resist AI Companions," preprint, May 21, http://dx.doi.org/10.2139/ssrn.5097445

De Freitas, Julian, Ahmet Kaan Uğuralp, Zeliha Oğuz-Uğuralp, and Stefano Puntoni (2024b), "Chatbots and Mental Health: Insights into the Safety of Generative AI," *Journal of Consumer Psychology*, 34 (3), 481–91.

De Gennaro, Mauro, Eva G. Krumhuber, and Gale Lucas (2020), "Effectiveness of an Empathic Chatbot in Combating Adverse Effects of Social Exclusion on Mood," *Frontiers in Psychology*, 10, 3061.

Devlin, Jacob, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova (2019), "Bert: Pre-Training of Deep Bidirectional Transformers for Language Understanding," in *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long And Short Papers)*, Minneapolis, MN: Association for Computational Linguistics, 4171–86.

Eccles, Alice M. and Pamela Qualter (2021), "Alleviating Loneliness in Young People: A Meta-Analysis of Interventions," *Child and Adolescent Mental Health*, 26 (1), 17–33.

Epley, Nicholas and Juliana Schroeder (2014), "Mistakenly Seeking Solitude," *Journal of Experimental Psychology. General*, 143 (5), 1980–99.

Esch, Franz-Rudolf, Tobias Langner, Bernd H. Schmitt, and Patrick Geus (2006), "Are Brands Forever? How Brand Knowledge and Relationships Affect Current and Future Purchases," *Journal of Product & Brand Management*, 15 (2), 98–105.

Fournier, Susan (1998), "Consumers and Their Brands: Developing Relationship Theory in Consumer Research," *Journal of Consumer Research*, 24 (4), 343–53.

Gable, Shelly L. and Harry T. Reis (2010), "Good News! Capitalizing on Positive Events in an Interpersonal Context," in *Advances in Experimental Social Psychology*, Vol. 42, ed. M. P. Zanna, San Diego, CA; and London: Elsevier, 195–257.

Gasteiger, Norina, Kate Loveys, Mikaela Law, and Elizabeth Broadbent (2021), "Friends from the Future: A Scoping Review of Research into Robots and Computer Agents to Combat Loneliness in Older People," *Clinical Interventions in Aging*, 16, 941–71.

Hayes, Andrew F. (2012), "Process: A Versatile Computational Tool for Observed Variable Mediation, Moderation, and Conditional Process Modeling [White Paper]," Last Accessed February 1, 2024. http://www.afhayes.com/public/process2012.pdf.

Henderson, Claire, Jo Noblett, Hannah Parke, et al. (2014), "Mental Health-Related Stigma in Health Care and Mental Health-Care Settings," *The Lancet Psychiatry*, 1 (6), 467–82.

Henseler, Jörg, Christian M. Ringle, and Marko Sarstedt (2015), "A New Criterion for Assessing Discriminant Validity in Variance-Based Structural Equation Modeling," *Journal of the Academy of Marketing Science*, 43 (1), 115–35.

Hoffman, Donna L. and Thomas P. Novak (1996), "Marketing in Hypermedia Computer-Mediated Environments: Conceptual Foundations," *Journal of Marketing*, 60 (3), 50–68.

Holt-Lunstad, Julianne (2021), "Loneliness and Social Isolation as Risk Factors: The Power of Social Connection in Prevention," *American Journal of Lifestyle Medicine*, 15 (5), 567–73.

Holt-Lunstad, Julianne, Theodore F. Robles, and David A. Sbarra (2017), "Advancing Social Connection as a Public Health Priority in the United States," *The American Psychologist*, 72 (6), 517–30.

Holzwarth, Martin, Chris Janiszewski, and Marcus M. Neumann (2006), "The Influence of Avatars on Online Consumer Shopping Behavior," *Journal of Marketing*, 70 (4), 19–36.

Hughes, Mary Elizabeth, Linda J. Waite, Louise C. Hawkley, and John T. Cacioppo (2004), "A Short Scale for Measuring Loneliness in Large Surveys: Results From Two Population-Based Studies," *Research on Aging*, 26 (6), 655–72.

Itzchakov, Guy, Netta Weinstein, Dvori Saluk, and Moty Amar (2023), "Connection Heals Wounds: Feeling Listened to Reduces Speakers' Loneliness Following a Social Rejection Disclosure," *Personality & Social Psychology Bulletin*, 49 (8), 1273–94.

Jiménez-Barreto, Jano, Natalia Rubio, and Sebastian Molinillo (2023), "How Chatbot Language Shapes Consumer Perceptions: The Role of Concreteness and Shared Competence," *Journal of Interactive Marketing*, 58 (4), 380–99.

Kahlon, Maninder K., Nazan Aksan, Rhonda Aubrey, et al. (2021), "Effect of Layperson-Delivered, Empathy-Focused Program of Telephone Calls on Loneliness, Depression, and Anxiety among Adults during the Covid-19 Pandemic: A Randomized Clinical Trial," *JAMA Psychiatry*, 78 (6), 616–22.

Kardas, Michael, Amit Kumar, and Nicholas Epley (2022), "Overly Shallow?: Miscalibrated Expectations Create a Barrier to Deeper Conversation," *Journal of Personality and Social Psychology*, 122 (3), 367–98.

Kull, Alexander J., Marisabel Romero, and Lisa Monahan (2021), "How May I Help You? Driving Brand Engagement through the Warmth of an Initial Chatbot Message," *Journal of Business Research*, 135, 840–50.

Leite, Iolanda, André Pereira, Samuel Mascarenhas, Carlos Martinho, Rui Prada, and Ana Paiva (2013), "The Influence of Empathy in Human–Robot Relations," *International Journal of Human–Computer Studies*, 71 (3), 250–60.

Lim, Eric (2020), "App-Store-Scraper." Last Accessed February 1, 2024. https://pypi.org/project/app-store-scraper/.
Liu, Lijun, Zhenggang Gou, and Junnan Zuo (2016), "Social Support Mediates Loneliness and Depression in Elderly People," *Journal of Health Psychology*, 21 (5), 750–8.
Liu, Yinhan, Myle Ott, Naman Goyal, et al. (2019), "Roberta: A Robustly Optimized Bert Pretraining Approach," arXiv, preprint, arXiv:1907.11692, July 26, https://doi.org/10.48550/arXiv.1907.11692
Liu-Thompkins, Yuping, Shintaro Okazaki, and Hairong Li (2022), "Artificial Empathy in Marketing Interactions: Bridging the Human-AI Gap in Affective and Social Customer Experience," *Journal of the Academy of Marketing Science*, 50 (6), 1198–218.
Longoni, Chiara, Andrea Bonezzi, and Carey K. Morewedge (2019), "Resistance to Medical Artificial Intelligence," *Journal of Consumer Research*, 46 (4), 629–50.
Lopez, Antonella, Alessandro O. Caffò, Luigi Tinella, and Andrea Bosco (2023), "The Four Factors of Mind Wandering Questionnaire: Content, Construct, and Clinical Validity," *Assessment*, 30 (2), 433–47.
Luo, Xueming, Marco Shaojun Qin, Zheng Fang, and Zhe Qu (2021), "Artificial Intelligence Coaches for Sales Agents: Caveats and Solutions," *Journal of Marketing*, 85 (2), 14–32.
Maples, Bethanie, 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.
Markov, Todor, Chong Zhang, Sandhini Agarwal, et al. (2023), "A Holistic Approach to Undesired Content Detection in the Real World," in *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 37, Washington, DC: AAAI (Association for the Advancement of Artificial Intelligence), 15009–18.
Marriott, Hannah R. and Valentina Pitardi (2024), "One Is the Loneliest Number... Two Can Be as Bad as One. The Influence of AI Friendship Apps on Users' Well-Being and Addiction," *Psychology & Marketing*, 41 (1), 86–101.
Masi, Christopher M., Hsi-Yuan Chen, Louise C. Hawkley, and John T. Cacioppo (2011), "A Meta-Analysis of Interventions to Reduce Loneliness," *Personality and Social Psychology Review: An Official Journal of the Society for Personality and Social Psychology, Inc*, 15 (3), 219–66.
Merrill, Kelly Jr., Jihyun Kim, and Chad Collins (2022), "AI Companions for Lonely Individuals and the Role of Social Presence," *Communication Research Reports*, 39 (2), 93–103.
Montoya, Amanda K. and Andrew F. Hayes (2017), "Two-Condition Within-Participant Statistical Mediation Analysis: A Path-Analytic Framework," *Psychological Methods*, 22 (1), 6–27.
Muniz, Albert M. Jr. and Thomas C. O'Guinn (2001), "Brand Community," *Journal of Consumer Research*, 27 (4), 412–32.
Myers, Sharon (2000), "Empathic Listening: Reports on the Experience of Being Heard," *Journal of Humanistic Psychology*, 40 (2), 148–73.
Nass, Clifford and Youngme Moon (2000), "Machines and Mindlessness: Social Responses to Computers," *Journal of Social Issues*, 56 (1), 81–103.
Palgi, Yuval, Amit Shrira, Lia Ring, et al. (2020), "The Loneliness Pandemic: Loneliness and Other Concomitants of Depression, Anxiety and Their Comorbidity during the COVID-19 Outbreak," *Journal of Affective Disorders*, 275, 109–11.
Perlman, Daniel and Letitia Anne Peplau (1982), "Theoretical Approaches to Loneliness," *Loneliness: A Sourcebook of Current Theory, Research and Therapy*, 36, 123–34.
Poscia, Andrea, Jovana Stojanovic, Daniele Ignazio La Milia, et al. (2018), "Interventions Targeting Loneliness and Social Isolation among the Older People: An Update Systematic Review," *Experimental Gerontology*, 102, 133–44.
Provoost, Simon, Ho Ming Lau, Jeroen Ruwaard, and Heleen Riper (2017), "Embodied Conversational Agents in Clinical Psychology: A Scoping Review," *Journal of Medical Internet Research*, 19 (5), e151.
Reis, Harry T., Edward P. Lemay Jr, and , and Catrin Finkenauer (2017), "Toward Understanding Understanding: The Importance of Feeling Understood in Relationships," *Social and Personality Psychology Compass*, 11 (3), e12308.
Replika (2025), "Is Replika Sentient?," Last Accessed February 1, 2024. https://help.replika.com/hc/en-us/articles/360058852132-Is-Replika-sentient.
Roos, Carla Anne, Tom Postmes, and Namkje Koudenburg (2023), "Feeling Heard: Operationalizing a Key Concept for Social Relations," *PloS One*, 18 (11), e0292865.
Schaaff, Kristina, Caroline Reinig, and Tim Schlippe (2023), "Exploring Chatgpt's Empathic Abilities," in *2023 11th International Conference on Affective Computing and Intelligent Interaction (ACII)*, Cambridge, MA: IEEE, 1–8.
Shrum, L. J., Elena Fumagalli, and Tina M. Lowrey (2023), "Coping with Loneliness through Consumption," *Journal of Consumer Psychology*, 33 (2), 441–65.
Shum, Heung-Yeung, Xiao-dong He, and Di Li (2018), "From Eliza to Xiaoice: Challenges and Opportunities with Social Chatbots," *Frontiers of Information Technology & Electronic Engineering*, 19 (1), 10–26.
Solano, Cecilia H., Phillip G. Batten, and Elizabeth A. Parish (1982), "Loneliness and Patterns of Self-Disclosure," *Journal of Personality and Social Psychology*, 43 (3), 524–31.
Ta, Vivian, Caroline Griffith, Carolynn Boatfield, et al. (2020), "User Experiences of Social Support from Companion Chatbots in Everyday Contexts: Thematic Analysis," *Journal of Medical Internet Research*, 22 (3), e16235.
Tillmann-Healy, Lisa M. (2003), "Friendship as Method," *Qualitative Inquiry*, 9 (5), 729–49.
Valenzuela, Ana, Stefano Puntoni, Donna Hoffman, et al. (2024), "How Artificial Intelligence Constrains the Human Experience," *Journal of the Association for Consumer Research*, 9 (3), 241–56.
Veronese, Nicola, Daiana Galvano, Francesca D'Antiga, et al. (2021), "Interventions for Reducing Loneliness: An Umbrella Review of Intervention Studies," *Health & Social Care in the Community*, 29 (5), e89–e96.
Welivita, Anuradha and Pearl Pu (2024), "Is Chatgpt More Empathetic Than Humans?," arXiv, preprint, arXiv:2403.05572, February 22, https://doi.org/10.48550/arXiv.2403.05572
Wilson, Timothy D. and Daniel T. Gilbert (2003), "Affective Forecasting," *Advances in Experimental Social Psychology*, 35 (35), 345–411.
Zielinski, Melissa J. and Jennifer C. Veilleux (2018), "The Perceived Invalidation of Emotion Scale (Pies): Development and Psychometric Properties of a Novel Measure of Current Emotion Invalidation," *Psychological Assessment*, 30 (11), 1454–67.

Journal of Consumer Research, 2025, 00, 1–23
https://doi.org/10.1093/jcr/ucaf040
Article