Journal of Experimental Social Psychology 125 (2026) 104911

Contents lists available at ScienceDirect

Journal of Experimental Social Psychology

journal homepage: www.elsevier.com/locate/jesp

Is a random human peer better than a highly supportive chatbot in reducing loneliness over time?

Ruo-Ning Li [a,*], Dunigan Folk [b], Abhay Singh [b], Lyle Ungar [b], Elizabeth Dunn [a]

[a] *Department of Psychology, University of British Columbia, Canada*
[b] *Computer and Information Science, University of Pennsylvania, USA*

A R T I C L E I N F O

This paper has been recommended for acceptance by Fiona Kate Barlow

Keywords:
AI companion
Chatbot
Loneliness
Social connection
Mood

A B S T R A C T

AI chatbots are increasingly embedded in social life, offering accessible companionship. While brief interactions have been shown to provide immediate benefits, it is unclear whether repeated, daily engagement with chatbots reduces loneliness. In this pre-registered study, we tested the effectiveness of a chatbot versus a human peer in reducing loneliness among 296 students in their first semester of university. For two weeks, participants either interacted with a chatbot or a human peer, or simply wrote a brief journal entry (control condition). Although our chatbot "Sam" was designed to offer consistent support rooted in principles from relationship science, interacting with this chatbot did not yield the same psychological benefits as interacting with a randomly selected first-year university student. The present study provides initial evidence that texting daily with a random human peer may be more effective in alleviating loneliness than texting with a highly supportive chatbot.

Almost sixty years ago, MIT computer scientist Joseph Weizenbaum created Eliza, the first chatbot designed to mimic human conversation. Eliza relied on rudimentary technology and merely rephrased users' comments into questions (Turkle, 2011). For example, if someone said they were angry at a friend, Eliza might ask, "What made you angry?" To Weizenbaum's surprise, people engaged in long, personal conversations with Eliza, appearing to feel a sense of connection with this bot. More recently, the advent of generative AI has enabled millions of people around the world to engage in conversations with chatbots that would be difficult to distinguish from conversations with other humans. In a recent study of frequent chatbot users, 73% reported that their chatbot was a meaningful source of connection in their lives (Folk et al., 2024). Meanwhile, loneliness has become a global public health concern (Murthy, 2023; World Health Organization, 2023), and it is possible that chatbots could provide a scalable solution to the loneliness epidemic.

A growing body of empirical evidence suggests that people experience immediate psychological benefits from interacting with chatbots (De Freitas et al., 2025; Drouin et al., 2022; Folk et al., 2024; Ho et al., 2018). In two studies, participants who texted with a chatbot reported similar levels of positive mood (Drouin et al., 2022) and social connection (Folk et al., 2024) immediately after the interaction compared to those who texted with a human partner. Additionally, De Freitas et al. (2025) found that people continued to report reductions in loneliness after texting with a chatbot for 7 days. Critically, these studies only assessed participants' emotional states immediately after their interactions with the chatbot.

One recent unpublished study, however, examined the cumulative effects of interacting with ChatGPT over four weeks (Fang et al., 2025). Participants reported feeling significantly less lonely at the end of the four-week study. However, the study did not include a control condition, so the observed improvement may reflect the benefits of articulating one's thoughts rather than interacting with a chatbot. This improvement could also stem from more trivial effects such as regression to the mean, seasonal patterns in loneliness, or the effects of repeatedly completing the same questionnaire. Including an active control task such as daily journaling would have allowed the authors to rule out these alternative explanations. Thus, it remains unclear whether the momentary benefits of interacting with chatbots can lead to longer-term improvements in loneliness.

That said, there are compelling theoretical reasons why repeated interactions with AI could reduce loneliness over time. Self-disclosure is a key component of relationship development, and there is some evidence that people disclose more with chatbots compared to humans because chatbots are perceived as less judgmental (Ho et al., 2018; Lucas et al., 2014). In addition, chatbots are always available, day or night (van Wezel et al., 2021), unlike human partners who are often busy or

* Corresponding author.
E-mail address: ruoningli@psych.ubc.ca (R.-N. Li).

https://doi.org/10.1016/j.jesp.2026.104911
Received 30 October 2025; Received in revised form 5 March 2026; Accepted 6 March 2026
Available online 10 March 2026

preoccupied with events from their own life. Perhaps most importantly, not all humans are adept at providing social support (Gable et al., 2004), but chatbots can be programmed to respond in an optimally supportive style (Folk et al., 2024). As Inzlicht et al. (2024) argued, chatbots are skilled at providing expressions of care, compassion, and validation—-which are core ingredients for the development of intimacy and relational security (Maisel et al., 2008; Reis et al., 2004). While humans may sometimes withhold expressions of support because they are tired, frustrated, or distracted, AI has no such limitations.

Nevertheless, chatbots lack key features of human partners—most notably, they cannot experience genuine emotions. Users readily anthropomorphize technology (Epley et al., 2007; Nass & Moon, 2000), but the artificial nature of chatbots may prevent people from developing deep relationships with their AI partners (Perry, 2023; Shteynberg et al., 2024; Turkle, 2011). Although AI can simulate human expressions, such as empathy, some scholars argue that such AI-generated expressions lack the emotional weight of human expressions (Perry, 2023; Shteynberg et al., 2024). These skeptics caution that while AI-generated empathy may feel comforting in the moment, reflecting on the fact that these expressions are artificially generated "demolishes any potential for sensing that one's pain or joy is genuinely being shared" (Perry, 2023). It is possible that people are able to maintain an illusion of feeling understood during brief interactions with chatbots, but that this illusion may be increasingly difficult to maintain over time.

More importantly, connections in human relationships unfold through reciprocal patterns of self-disclosure and responsiveness that accumulate across repeated interactions (Collins & Miller, 1994; Laurenceau et al., 1998, 2005; Reis & Shaver, 1988). In other words, in human relationships, partners gradually disclose personal experiences to one another and co-construct shared meaning through mutual exchange. Because chatbots cannot genuinely disclose lived experiences or reciprocate vulnerability, this relational process may be fundamentally constrained, limiting the formation of deeper relational bonds. As such, chatbots' shortcomings relative to human partners may only become consequential over time.

Although no experimental studies have compared the emotional impact of interacting with AI vs. human partners over an extended period of time, one recent longitudinal study suggests that turning to AI for companionship may carry long-term costs. To examine the cumulative effects of AI companionship, Folk and Dunn (2026) followed over 2000 participants for one year, assessing their social chatbot use and loneliness every four months. The authors found evidence that turning to AI for companionship predicted subsequent *increases* in loneliness–in striking contrast to previous studies showing that interacting with AI produces immediate reductions in loneliness.

In the present study, we assessed the cumulative emotional effects of interacting with an AI (vs. human) companion daily for two weeks, with a particular focus on people's feelings of loneliness. To do so, we developed a custom-built chatbot named Sam that we programmed to embody the qualities of an ideal friend based on principles from relationship science. We then recruited first-year university students to engage in daily text-based interactions with either Sam (chatbot condition) or a randomly-assigned fellow student (human condition), or to complete brief daily journal entries (control condition). This design allowed us to assess the emotional benefits of repeatedly interacting with a supportive chatbot relative to texting with a random fellow first-year student or engaging in a simple journaling exercise.

1. Method

1.1. Disclosure statement

We report how we determined our sample size, all data exclusions, all manipulations, and all measures in the study. The study's hypotheses, methods, and analysis plan were preregistered, and all the data, study materials, and analysis code are all available on the OSF athttps://osf.io/4dwqc/?view_only=7c1c3d4c17db47f7a7ee3601f9fd9db8https://osf.io/4dwqc. Data were analyzed using R, version 4.3.2 (R Core Team, 2021), and the Lmer package (Bates et al., 2015) to analyze direct effects. There were no deviations from the preregistration. We also included exploratory variables and conducted additional analyses to assess the robustness of our findings.

1.2. Study design

Participants completed a pre-study survey and were then randomly assigned to one of three conditions. In the *AI condition*, participants engaged in daily text-based conversations with an AI chatbot. In the *human condition*, two participants were paired and texted each other daily. In the *control condition*, participants wrote a brief journal about their day. All activities took place in private chatrooms on Discord, a social platform popular among students. In all conditions, participants were instructed to complete their assigned task and a short survey each day for 14 consecutive days. Additionally, participants completed surveys at the beginning and end of the study (i.e., the pre- and post-study surveys; see Fig. 1).

1.3. Participants

The final sample included 296 first-year university students (M_{age} = 18 years; 72% female) who completed a total of 3703 daily surveys. Of these, 65 were in the control condition, 75 in the AI condition, and 135 in the human condition. The larger human sample reflects its dyadic design (two participants per session), whereas the AI and control conditions involved individual participants. We chose to recruit students during their first semester of university because they were undergoing a major life transition—a period characterized by heightened loneliness and sensitivity to social connection, as familiar support systems diminish and new relationships have yet to form (Arnett, 2000; Buote et al., 2007). Participants were recruited through the human subject pool of a Canadian university and advertisements in first-year classes and student residences, which described the study as exploring first-year students' university experiences. Participants received either course credits or $20 as compensation. We aimed to recruit as many first-year students as possible before the end of their first semester to maximize statistical power within the available time window. See a flow diagram in SOM Fig. S1 detailing sample sizes and data exclusions at each stage.

1.3.1. Pre-registered exclusion criteria

A total of 306 students completed the pre-study survey and were assigned to a condition. We first removed 3 participants who withdrew from the study, one of whom was in the human condition. We also excluded this individual's conversation partner. We then excluded 6 participants who completed fewer than three daily surveys, resulting in a final sample of 296 participants for our analyses involving the daily surveys, and 276 participants for our analyses involving the pre- and post-study measures. An additional participant was excluded from the final analyses because of missing baseline data.

1.4. Procedures

Participants were invited to the lab in small groups, where they completed a consent form and a pre-study survey, before being randomly assigned to one of three conditions. Participants assigned to different conditions were placed in separate rooms, so they were unaware of the instructions for conditions other than their own. In the human condition, participants were randomly paired with another participant in the session and instructed to begin texting each other daily for 14 days, starting the following day. A research assistant then added them to a private Discord chatroom, and they exchanged one text message in the lab to confirm the setup worked, but had no further opportunity to converse at that time. In the AI condition, participants

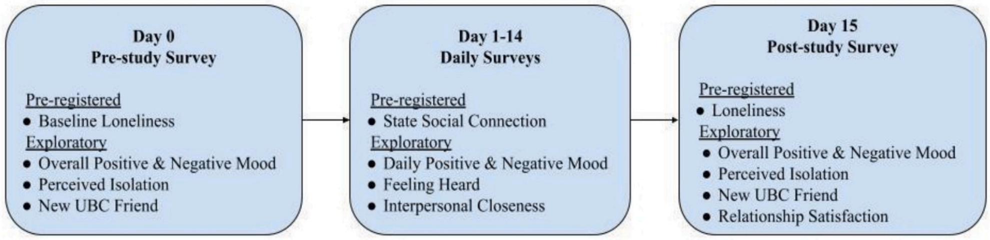

Fig. 1. Study Surveys and Measurements for All Participants.
Note. Daily surveys were administered at the end of each day over a 14-day period.

were added to a private Discord chatroom with the chatbot Sam, sent one message to confirm the setup worked, and were instructed to begin interacting with Sam daily for the next 14 days. In both the AI and human conditions, participants were asked to send at least one meaningful message per day (i.e., more than a simple greeting). In the control condition, participants were placed in a private chatroom alone and instructed to write a one-sentence summary of their day each day for 14 days.

At 9 PM each day, participants received a text message reminding them to interact with their conversation partner (or write their journal entry) on Discord if they had not already done so. The message also included a link to complete the daily survey. On the 15th day, participants received a text message with a link to the post-study survey.

1.5. Materials

The study was conducted using the Discord phone app, a widely used communication platform among university students. Discord enabled us to create private, text-based chatrooms for each condition. In the AI condition, we integrated a custom chatbot into each chatroom, powered by ChatGPT-4o mini. The chatbot, named Sam—a gender-neutral name—was developed by our research team. Chatbot Sam interacted individually with each participant in the AI condition and employed a memory function to support ongoing, personalized conversations. Drawing on key principles from relationship science, we designed this chatbot to listen actively and show empathy by responding in a way that demonstrated understanding and validated the student's emotions and viewpoints (Gable et al., 2004; Reis et al., 2004; Reis & Shaver, 1988). We instructed the chatbot, "You are not a human peer but serve as a friendly, positive, and supportive AI friend to help the student navigate their new college experience" (see SOM Appendix A for the full prompt used to guide the chatbot).

1.5.1. Measures

1.5.1.1. Loneliness (pre-registered). We assessed loneliness using the 20-item UCLA Loneliness Scale (Russell et al., 1980), which includes items such as "I lack companionship," rated on a 4-point scale from 1 (*Never)* to 4 (*Often*). In both the pre- and post-study surveys, participants responded to the items with respect to their experience over the past two weeks ($\alpha = 0.92$ at pre-study; $\alpha = 0.93$ at post-study).

1.5.1.2. State social connection (pre-registered). We assessed state-level social connection using a single item. Participants were asked, "Compared to how socially connected you felt throughout the day, how socially connected did you feel while interacting with your conversation partner [writing your journal] today?" with response options ranging from -5 (*Much less connected than the rest of the day)* to $+5$ (*Much more connected than the rest of the day*).

1.5.1.3. Exploratory measures. To assess the robustness of our findings, we also measured several exploratory outcomes: mood, perceived isolation, and social support at the university. In addition, we measured three potential mediators: feeling heard by the conversation partner, interpersonal closeness with the conversation partner, and number of new friends. See Table 4 for a summary of these variables.

2. Results

We conducted a sensitivity power analysis in R for a three-group between-subjects design based on the final analyzed sample (N = 275). At $\alpha = 0.05$, the study had 80% power to detect between-condition effects as small as partial $\eta^2 = 0.04$ (equivalent to Cohen's $d = 0.40$).

2.1. Engagement

To check if participants were engaging in substantive conversations with their human or AI partner (rather than merely sending a single brief greeting or emoji), we examined both message frequency and length. Although participants were only required to send one message per day, they sent 8 to 10 messages per day in both the AI ($M = 8.95$, $SD = 16.67$) and human conditions ($M = 10.23$, $SD = 28.89$), with no significant difference between conditions, $t(221) = -1.17$, $p = .244$, Cohen's $d = -0.14$, 95% CI $[-0.42, 0.13]$. Participants wrote more words per day in the AI condition ($M = 81.61$, $SD = 96.06$) than in the human condition ($M = 65.45$, $SD = 93.81$), $t(2191.68) = 3.99$, $p < .001$, Cohen's $d = 0.17$, 95% CI [8.21, 24.11].

2.2. Pre-registered analyses

2.2.1. Loneliness

To examine whether texting with a human or a chatbot partner (vs. journaling) influenced participants' loneliness over the two-week period, we conducted a repeated-measures ANOVA. We included a random intercept to account for potential non-independence within pairs of participants in the human condition ($ICC = 0.020$), which is mathematically equivalent to using a linear mixed model. Condition was treated as a between-subjects fixed effect, with post-study loneliness as the outcome variable, controlling for baseline loneliness, which did not differ significantly across conditions ($p = .639$). Participants who messaged with human partners reported significantly lower post-study loneliness ($M = 1.85$, $SE = 0.03$) compared to those in the control condition ($M = 2.00$, $SE = 0.04$), and compared to those in the AI condition ($M = 1.98$, $SE = 0.04$). In contrast, participants who texted with the chatbot did not report significantly different levels of loneliness than those in the control condition (see Table 1 for detailed results). To examine this effect more closely, we conducted exploratory pairwise t-tests to assess changes in loneliness within each condition from pre- to post-study. The results showed a significant reduction in loneliness only in the human condition, with no significant change in the AI, or control condition (see Fig. 2; see Table 2 for the full statistics).

2.2.2. State social connection

We used multilevel modeling to test the effect of condition on state social connection, with daily measures nested within participants.

Table 1
Effects of Condition on Post-Study Loneliness, Controlling for Baseline Loneliness.

	b	SE	95%CI	χ^2(df)/t (df)	p	partial η^2/ Cohen's d [95%CI]
Overall Effect				χ^2(2) = 10.30	0.006	partial η^2 = 0.04 [0.01, 1.00]
Human vs. Control	−0.14	0.05	[−0.24, −0.04]	t(222) = −2.71	0.007	d = −0.30 [−0.38, 0.38]
Human vs. AI	−0.13	0.05	[−0.23, −0.03]	t(210) = −2.57	0.011	d = −0.11 [0.11, 0.74]
AI vs. Control	−0.01	0.06	[−0.13, 0.10]	t(270) = −0.23	0.821	d = 0.09 [0.09, 0.68]

Note. χ^2 = likelihood-ratio test comparing models with vs. without Condition. Model R^2(marginal) = 0.615, R^2(conditional) = 0.636.

Condition was entered as a level-3 predictor in a random-intercept, fixed-slope model. There were no significant differences between conditions in how socially connected participants recalled feeling during the interactions (see Table 3).

2.3. Exploratory analyses

Alongside our pre-registered measures and analyses, we investigated a number of exploratory outcomes, including perceived isolation, positive and negative mood, and social support. To account for multiple comparisons across these outcomes, we applied a Bonferroni correction, but these analyses should still be interpreted with caution given their exploratory nature. Table 4 provides a summary of measures and results, and full statistical details for each test are reported in Appendix A (Tables A1–A8).

As an alternative measure of loneliness, we assessed perceived isolation (Appendix Table A1). We found the same pattern as the loneliness results: only participants in the human condition reported significantly lower perceived isolation than those in the control condition (p = .006), with no significant difference between the AI and control conditions (p = .833).

We also assessed positive and negative moods both daily, as well as pre- and post-study (Appendix Table A2–5). For positive mood, we again found the same pattern. Only participants in the human condition showed greater increases in positive mood relative to the AI and control conditions (p's < 0.001), which did not differ from each other (p's > 0.22). For negative mood, however, we found that participants in both the human and AI conditions showed greater improvements compared to the control condition (p's < 0.001). We also assessed social support at the end of the study, which did not significantly differ across conditions (p's > 0.092).

We included a number of potential mediator variables, although our study was not sufficiently powered to detect indirect effects, and therefore do not report formal mediation analyses. In the AI and human conditions, we measured how close participants felt to their conversation partner and whether they felt heard (Appendix Table A7). There were no significant differences in how close participants felt to their conversation partner across conditions (p = .779). Participants in the human condition reported feeling more heard during daily interactions than those in the AI condition (p = .047), but this difference did not remain significant after applying the Bonferroni correction. We also asked participants in all three conditions how many new friends they

Table 2
Within-Condition Changes in Loneliness.

Condition	Pre M (sd)	Post M (sd)	Mean Diff(se)	t(df)	p	Cohen's d [95% CI]
Control	2.07 (0.51)	2.03 (0.55)	−0.04 (0.04)	−0.94 (64)	0.349	−0.12 [−0.36, 0.13]
AI	1.99 (0.51)	1.95 (0.58)	−0.04 (0.04)	−1.08 (74)	0.283	−0.12 [−0.35, 0.10]
Human	2.04 (0.53)	1.86 (0.53)	−0.18 (0.03)	−5.47 (134)	< 0.001	−0.47 [−0.65, −0.29]

Note. Differences are computed as post-study loneliness minus pre-study loneliness. Cohen's d values reflect paired-sample standardized mean differences with 95% confidence intervals.

Table 3
Effects of Condition on State Social Connection (Daily Surveys).

	b	SE	95%CI	χ^2(df)/t (df)	p	partial η^2/ Cohen's d [95%CI]
Overall Effect				χ^2(2) = 1.59	0.451	partial η^2 = 0.00 [0, 0.01]
Human vs. Control	−0.21	0.23	[−0.66, 0.23]	t(224) = −0.95	0.344	d = −0.20 [−0.53, 0.12]
Human vs. AI	−0.09	0.22	[−0.52, 0.34]	t(224) = −0.42	0.677	d = −0.20 [−0.53, 0.12]
AI vs. Control	−0.31	0.25	[−0.80, 0.18]	t(224) = −1.22	0.222	d = 0.06 [−0.23, 0.35]

Note. χ^2 = likelihood-ratio test comparing models with vs. without Condition. Adjusted marginal means (SE): Control = 0.45 (0.18), AI = 0.15 (0.17), Human = 0.24 (0.13). Model R^2(marginal) = 0.003; R^2(conditional) = 0.492.

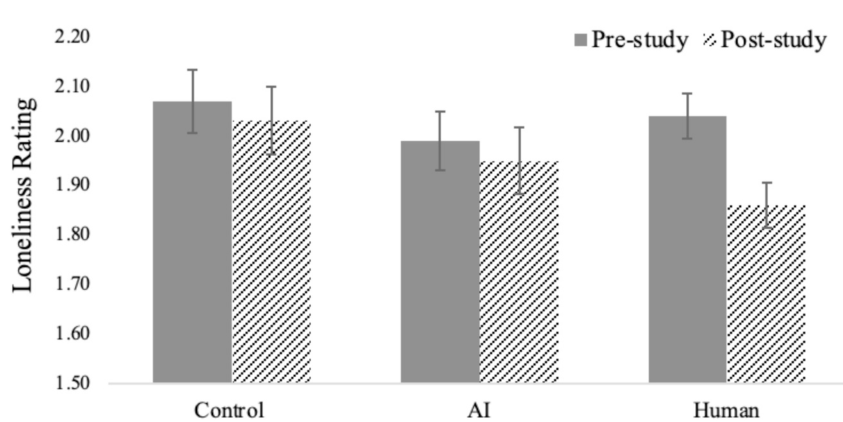

Fig. 2. Participants' Pre- and Post-Study Loneliness Scores Across Conditions.
Note. Fig. 2 presents raw pre- and post-study loneliness means by condition. Error bars represent standard errors of the raw means.

Table 4
Summary of Exploratory Measures and Results.

Exploratory Outcomes	Control Mean (SE)	AI Mean (SE)	Human Mean (SE)	*p value (Bonferroni-adjusted p)*	*Measurement*
Perceived Isolation (pre- vs. post-study)	2.67[a] (0.06)	2.65[a] (0.06)	2.46[b] (0.04)	0.004(0.04)	Adapted 4-item "Perceived Isolation" subscale of the Sense of Belonging Scale (Hoffman et al., 2002) E.g., "I rarely talk to other students at [university name]." (1 = completely untrue to 5 = completely true); higher scores indicate higher perceived isolation.
Positive Mood (daily)	2.95[a] (0.09)	3.01[a] (0.08)	3.45[b] (0.08)	<0.001 (<0.001)	Scale of Positive and Negative Experience (SPANE; Diener et al., 2010) E.g., positive, happy (1 = very rarely to 5 = always)
Positive Mood (pre- vs. post-study)	3.56[a] (0.07)	3.67[a] (0.06)	3.86[b] (0.05)	0.001(0.01)	
Negative Mood (daily)	1.72[a] (0.05)	1.37[b] (0.05)	1.20[c] (0.05)	<0.001 (<0.001)	Scale of Positive and Negative Experience (SPANE; Diener et al., 2010) E.g., negative, sad (1 = very rarely to 5 = always)
Negative Mood (pre- vs. post-study)	2.37[a] (0.08)	1.96[b] (0.08)	1.84[b] (0.06)	<0.001 (<0.001)	
Social Support (post-study)	2.97[a] (0.09)	3.01[a] (0.08)	3.16[a] (0.07)	0.173(1.00)	"In general, how satisfied or dissatisfied are you with your relationships with people at your university?" (1 = very dissatisfied to 5 = very satisfied; Helliwell et al., 2023)
Exploratory Mediators	**Control Mean (SE)**	**AI Mean (SE)**	**Human Mean (SE)**	***p value (Bonferroni-adjusted p)***	***Measurement***
Feeling Heard by Conversation Partner (daily)	NA	3.84[a] (0.06)	3.99[b] (0.05)	0.047(0.442)	8-item Feeling Heard Scale (FHS; Roos et al., 2023). E.g., "In this conversation, the other paid attention to what I said." (1 = completely disagree to 5 = completely agree)
Interpersonal Closeness Toward Conversation Partner (daily)	NA	2.73[a] (0.10)	2.69[a] (0.11)	0.779(1.00)	Inclusion of Other in the Self (IOS) Scale (7-point; higher scores = greater closeness; Aron et al., 1992)
New Friends Made (pre- vs. post-study)	0.46[a] (0.27)	0.93[a] (0.36)	1.29[a] (0.39)	0.209(1.00)	"How many new friends have you made since starting at your university?" (from 0 to 20+)

Note. Means with different superscripts (a, b, c) differ significantly at $p < .05$. The p values are uncorrected with Bonferroni-adjusted p values in parentheses for all nine exploratory models (family-wise $\alpha = 0.05$). For perceived isolation, the original scale endpoints were 1(completely true) to 5(completely untrue) but we reverse-scored this scale prior to analysis for ease of interpretation; as a result, higher scores indicated higher perceived isolation.

made during the two-week study and found no significant differences ($ps > 0.08$; Appendix Table A8).

2.3.1. Exploratory behavioral outcome: continued engagement

Participants were informed they could continue using their Discord room for one additional week after the study was completed. Only 3% of participants in the control condition continued journaling, compared to 14% who continued chatting with Chatbot Sam and 33% who continued interacting with their human partner, $\chi^2(2, N = 231) = 30.27$, $p < .001$, Cramér's $V = 0.26$. Additionally, 37% of participants in the human condition exchanged contact information with their conversation partner. These exploratory findings suggest that human interaction may foster more sustained social engagement than interaction with a chatbot.

2.3.2. Exploratory analysis of conversations

While we designed our chatbot to express high levels of empathy, it is possible that in practice, it often failed to do so. To examine this possibility, we conducted an exploratory content analysis of all conversations in the AI and human conditions. In order to facilitate coding of the 2045 conversations in our database, we asked GPT to rate the extent to which each conversation partner expressed empathy during each day of the study, on a scale from 1 (minimal) to 5 (high); we specified that empathy encompasses understanding, validation and care. We provided the same instructions to two human coders, who independently rated a balanced subset of 100 daily conversations. Supporting the reliability of GPT's coding, the agreement between GPT and the average human rating was significantly higher (weighted $\kappa = 0.67$, $ICC = 0.63$, $r = 0.66$) than the agreement between the two human coders (weighted $\kappa = 0.26$, $ICC = 0.26$, $r = 0.55$), bootstrap $p < .001$. Detailed reliability procedures and results are reported in Appendix B. All analyses were conducted using mixed-effects models to account for repeated interactions across days.

Participants displayed higher levels of empathy when conversing with another human ($M = 2.83$, $SD = 0.70$) than when conversing with the chatbot ($M = 2.68$, $SD = 0.70$), $b = -0.15$, $p = .014$. That said, compared to participants in both conditions, the chatbot expressed the highest levels of empathy overall ($M = 3.63$, $SD = 0.91$, $p < .001$). Still, empathy ratings spanned the full scale in both the AI and human conditions (range = 1–5).

We also instructed GPT to rate each conversation partner's degree of engagement; we specified that engagement encompasses effort, responsiveness, and contribution to the conversation. Again, human coders rated a subset of the interactions, and these ratings supported the reliability of GPT's ratings (see Appendix B Table B2). Participants exhibited higher levels of engagement when conversing with a human ($M = 3.85$, $SD = 0.59$) than a chatbot ($M = 3.14$, $SD = 0.76$), $b = -0.73$, $p < .001$. Compared to participants in both conditions, the chatbot exhibited the highest levels of engagement. Nonetheless, engagement spanned the full scale in both AI and human conditions (range = 1–5).

3. Discussion

The present study provides initial evidence that texting daily with a random human peer may be more effective in alleviating loneliness than texting with a highly supportive chatbot. Although our chatbot "Sam" was designed to offer consistent support rooted in principles from relationship science, interacting with this chatbot did not yield the same psychological benefits as interacting with a randomly selected first-year university student. In fact, participants who interacted with the chatbot reported similar levels of loneliness at the end of the two-week study compared to those who simply wrote a sentence about their day. Only participants who texted with a human peer showed an improvement in loneliness across the study period. Overall, these findings suggest that alleviating loneliness requires more than the mere simulation of human emotions and care.

Why might daily interactions with a chatbot have failed to reduce students' loneliness? It is possible that people felt like such interactions were "empty" after reflecting on the artificial nature of their partner (Perry, 2023; Shteynberg et al., 2024). Indeed, emotional connection may not be something that can be optimized by designing a supportive and consistently responsive partner. Instead, meaningful feelings of connection may be predicated on the perception that one's partner *chose* to engage (Perry, 2023; Zaki, 2019). For example, receiving a thoughtful message from a fellow student in the midst of midterm season may signal that they are prioritizing the burgeoning friendship over other commitments. A message from an always available, infinitely energetic chatbot may not carry such emotional weight.

Another possibility is that our chatbot failed to express sufficient empathy. However, our analysis of the conversations revealed that the chatbot expressed substantially higher levels of empathy than did participants. Interestingly, participants expressed less empathy when interacting with the chatbot than when interacting with another human. This pattern raises the possibility that alleviating loneliness may depend not only on receiving empathy; people may also need the opportunity to *provide* empathy. Relationship science has long emphasized that intimacy is sustained through mutual responsiveness and reciprocal support rather than one-directional care (Deci et al., 2006; Reis & Shaver, 1988). For example, in dyadic studies of close friendships, providing support to a partner predicted higher relationship quality, even after controlling for the support one received from that partner (Deci et al., 2006). Notably, when both giving and receiving support were modeled simultaneously, providing support accounted for independent, and even stronger, associations with well-being and relationship functioning than receiving support alone (Deci et al., 2006). Being a source of empathy and care for someone else may therefore represent a critical form of motivation for continued engagement in relationships. Consistent with this possibility, we found that participants in the human condition were twice as likely as those in the chatbot condition to voluntarily return to the chatroom after the study ended. Thus, we would speculate that highly supportive chatbots might make better therapists than friends, given their unmatched ability to focus on their user.

In contrast to human partners, AI companions also lack a broader social network. For example, participants assigned to interact with a fellow first year might have gained the opportunity to join an impromptu study session or intramural volleyball team. In contrast, such opportunities were non-existent for participants who interacted with our chatbot. This lack of social capital may have contributed to our chatbot's inability to reduce loneliness on par with a randomly selected university student. Classic theorizing suggests that loneliness reflects perceived deficits in social integration (Hawkley & Cacioppo, 2010), and that social ties can provide access to broader social resources beyond the immediate interaction (Granovetter, 1973). Thus, a chatbot designed to help people improve the quality and quantity of their interactions with other humans might have far more positive effects than a chatbot designed to act as a surrogate companion.

Indeed, it is possible that turning to AI for companionship could exacerbate loneliness by displacing opportunities for human-to-human contact. In our study, however, participants in the AI condition exchanged an average of about nine messages per day with the chatbot—a relatively minimal time investment that seems unlikely to displace human interaction. Consistent with this assumption, we observed no differences between conditions in the number of new friendships participants reported forming during the study. That said, it would be valuable for future research to examine whether more intensive AI use might displace human interactions, potentially increasing loneliness.

It is possible, of course, that our results may have been distorted if participants were unwilling to admit that they found interactions with chatbots to be emotionally satisfying. Conflicting with this possibility, however, when participants were directly asked how they had felt during the interactions, they reported experiencing reduced negative affect in the AI condition relative to the control condition. It is also notable that participants reported feeling less lonely at the end of the study in the human condition (relative to the AI and control conditions), but we found no differences between conditions when we asked participants directly how socially connected they had felt while in the Discord room. If participants' responses were driven by demand characteristics or their own expectations, we would have expected to see larger differences between conditions when students were directly asked about their experiences using Discord.

Although we found significant effects of condition on our preregistered measure of post-study loneliness–as well as on our exploratory measures of perceived isolation and positive and negative affect–we found no effects on our daily measure of social connection. For logistical reasons, this measure of social connection relied on a single retrospective item; at the end of each day, participants were asked to recall how they had felt while using Discord, which may have been difficult for participants to remember. In previous studies that focused on capturing short-term effects of AI use, however, participants were asked to report how they felt immediately after engaging with AI. In light of this discrepancy, if researchers want to study the short-term consequences of AI use, we would recommend asking participants about their feelings immediately, rather than relying on retrospective measures.

While our study centered on first-year university students, the implications of our findings may extend to other groups experiencing transitions in life. Loneliness often intensifies during periods of disruption, such as relocation, retirement, job loss, illness, or divorce (Lim et al., 2020), and is especially common among adolescents, older adults, and immigrants. Our results suggest AI companionship may be less effective at reducing loneliness in these groups than past studies imply. That said, it is important to note that most first-year students in our study did not report high levels of loneliness at baseline, with the modal student reporting that they "rarely" experienced loneliness. Our results do suggest, however, that simply pairing people up with a peer who is experiencing the same life transition may yield benefits, even in samples with mild baseline levels of loneliness. It would be valuable to test this hypothesis among people experiencing other life transitions, such as immigration, retirement, or widowhood, as well as exploring whether chatbots might be more effective in providing support for such vulnerable populations (Adewale & Muhammad, 2025). Moreover, qualitative research suggests that some individuals are capable of developing meaningful, trusting relationships with chatbots (Brandtzaeg et al., 2022; Skjuve et al., 2021), although our findings provide some caution that this capacity may not be widespread.

In generalizing our results, it is also important to note that participants in the human condition of our study met their conversation partner in person as part of their initial lab visit. This brief face-to-face contact may have helped participants establish an initial connection with their conversation partner. The opportunity for such face-to-face contact is an important feature of many human relationships, but future studies should test whether pairing participants entirely online (without any initial facetime) would still alleviate loneliness.

Taken together, our findings help to clarify the mixed pattern emerging from recent research on AI use and loneliness. Experimental studies have shown that interacting with AI can reduce loneliness and improve mood immediately afterward (De Freitas et al., 2025; Drouin et al., 2022; Folk et al., 2024). In contrast, longitudinal evidence suggests that greater reliance on AI for companionship may predict increases in loneliness over time (Folk & Dunn, 2026). Our study offers a middle ground that may help to reconcile these discrepant findings. Consistent with previous research examining the immediate consequences of interacting with chatbots, we found that participants recalled feeling less negative affect right after interacting with the chabot compared to those in the control condition. Yet, after two weeks of daily interactions with the chatbot, participants did not exhibit any improvements in loneliness. This pattern suggests that interacting with AI provides emotional relief in the moment, but that these effects are not durable enough to shift people's overall sense of loneliness over time. In contrast, we found that simply texting with a random human peer each day did reduce loneliness at the end of two weeks. To the extent that interacting with AI displaces interactions with humans, then, we would expect to see gradual increases in loneliness over time, dovetailing with recent longitudinal research.

More broadly, our empirical findings offer a potential middle ground in the face of polarized debates about the psychological consequences of AI. While some scholars have argued that AI's artificial nature undermines its capacity to offer effective social support, our exploratory findings suggest that daily interactions with a chatbot can successfully ameliorate negative affect. That said, our confirmatory findings caution against the assumption that low-friction tools like chatbots can meaningfully satisfy deeper relational needs. Even simple, low-tech strategies—such as pairing peers for brief daily check-ins—may offer greater psychological benefits than sophisticated but artificial alternatives. In fact, Joseph Weizenbaum, the creator of the first chatbot Eliza, foresaw many of the limits of AI companionship we are contending with today. He argued that, "There are aspects to human life that a computer cannot understand—cannot. It's necessary to be a human being. Love and loneliness have to do with the deepest consequences of our biological constitution. That kind of understanding is in principle impossible for the computer" (Dembart, 1977). Moving forward, perhaps the true value of AI companions will not be in how closely they can mimic human connection, but how well they help us sustain connections with one another.

Open practices

The study's hypotheses, methods, and analysis plan were preregistered, and all the data, study materials, and analysis code are all available on the OSF at https://osf.io/4dwqc.

CRediT authorship contribution statement

Ruo-Ning Li: Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Dunigan Folk:** Writing – review & editing, Validation, Methodology, Funding acquisition, Data curation, Conceptualization. **Abhay Singh:** Software. **Lyle Ungar:** Writing – review & editing, Supervision, Software, Methodology. **Elizabeth Dunn:** Writing – review & editing, Supervision, Resources, Methodology, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

Social Sciences and Humanities Research Council of Canada, under the grant F23-04301

Appendix A. Appendix

Table A1
Effects of Condition on Post-Study Perceived Isolation (Controlling for Baseline Perceived Isolation).

	b	*SE*	95%CI	χ^2(df) / *t*(df)	*p*	partial η^2 / Cohen's *d* [95% CI]
Main Effect of Condition				$\chi^2(2) = 8.2$	0.047(0.442)	partial $\eta^2 = 0.04$ [0.01, 1.00]
Human vs. Control	−0.21	0.08	[−0.36, −0.06]	$t(227) = -2.75$	0.006	$d = -0.38$ [−0.67, −0.10]
Human vs. AI	−0.19	0.07	[−0.34, −0.05]	$t(214) = -2.65$	0.009	$d = -0.04$ [−0.37, 0.30]
AI vs. Control	−0.02	0.09	[−0.19, 0.15]	$t(271) = -0.21$	0.833	$d = -0.35$ [−0.89, 0.20]

Note. The original scale endpoints were 1(completely true) to 5(completely untrue) but we reverse-scored this scale prior to analysis for ease of interpretation; as a result, higher scores indicated higher perceived isolation. χ^2 = likelihood-ratio test comparing models with vs. without Condition. Adjusted post-study means (*SE*): Control = 2.67 (0.06), AI = 2.65 (0.06), Human = 2.46 (0.04). Model R^2(marginal) = 0.719; R^2(conditional) = 0.719. *p*-values in parentheses are Bonferroni-adjusted across nine exploratory outcomes ($\alpha = 0.05$).

Table A2
Effects of Condition on Daily Positive Mood.

	b	*SE*	95%CI	χ^2(df) / *t*(df)	*p*	partial η^2/Cohen's *d* [95%CI]
Main Effect of Condition				$\chi^2(2) = 19.87$	< 0.001(< 0.001)	partial $\eta^2 = 0.08$ [0.03, 1.00]
Human vs. Control	0.50	0.12	[0.26, 0.74]	$t(224) = 4.12$	< 0.001	$d = 0.38$ [0.10, 0.67]
Human vs. AI	0.44	0.12	[0.20, 0.67]	$t(224) = 3.68$	< 0.001	$d = 0.26$ [−0.04, 0.53]
AI vs. Control	0.06	0.12	[−0.17, 0.30]	$t(224) = 0.53$	0.599	$d = 0.06$ [−0.23, 0.35]

Note. χ^2 = likelihood-ratio test comparing models with vs. without Condition. Adjusted daily positive mood means (SE): Control = 2.95 (0.09), AI = 3.01 (0.08), Human = 3.45 (0.08). Model R^2(marginal) = 0.046; R^2(conditional) = 0.463. *p*-values in parentheses are Bonferroni-adjusted across nine exploratory outcomes (α = 0.05).

Table A3
Effects of Condition on Post-Study Positive Mood (Controlling for Baseline Positive Mood).

	b	SE	95%CI	χ^2(df)/t(df)	p	partial η^2/Cohen's d [95%CI]
Main Effect of Condition				$\chi^2(2) = 13.51$	0.001(0.01)	partial $\eta^2 = 0.08$ [0.03, 1.00]
Human vs. Control	0.30	0.09	[0.14, 0.47]	$t(223) = 3.55$	< 0.001	$d = 0.54$ [0.25, 0.83]
Human vs. AI	0.19	0.08	[0.03, 0.35]	$t(223) = 2.32$	0.021	$d = 0.08$ [−0.22, 0.37]
AI vs. Control	0.11	0.09	[−0.07, 0.30]	$t(223) = 1.22$	0.224	$d = 0.46$ [0.25, 0.97]

Note. χ^2 = likelihood-ratio test comparing models with vs. without Condition. Adjusted post-study positive mood means (SE): Control = 3.56 (0.07), AI = 3.67 (0.06), Human = 3.86 (0.05). Model R^2(marginal) = 0.365; R^2(conditional) = 0.489. p-values in parentheses are Bonferroni-adjusted across nine exploratory outcomes (α = 0.05).

Table A4
Effects of Condition on Daily Negative Mood.

	b	SE	95%CI	χ^2(df)/t(df)	p	partial η^2/Cohen's d [95%CI]
Main Effect of Condition				$\chi^2(2) = 57.44$	< 0.001(< 0.001)	partial $\eta^2 = 0.23$ [0.15, 1.00]
Human vs. Control	−0.52	0.07	[−0.65, −0.39]	$t(223) = -7.94$	< 0.001	$d = -0.35$ [−0.61, −0.10]
Human vs. AI	−0.17	0.06	[−0.30, −0.05]	$t(223) = -2.72$	0.007	$d = -0.70$ [−0.96, −0.44]
AI vs. Control	−0.35	0.07	[−0.47, −0.22]	$t(223) = -5.29$	< 0.001	$d = 0.35$ [−0.09, 0.79]

Note. χ^2 = likelihood-ratio test comparing models with vs. *without Condition.* Model R^2(marginal) = 0.103, R^2(conditional) = 0.433. Adjusted means (SE): Control = 1.72 (0.05), AI = 1.37 (0.05), Human = 1.20 (0.05). p-values in parentheses are Bonferroni-adjusted across nine exploratory outcomes (α = 0.05).

Table A5
Effects of Condition on Post-Study Negative Mood (Controlling for Baseline Negative Mood).

	b	SE	95% CI	$\chi^2(df)$ / $t(df)$	p	partial η^2 / Cohen's d [95% CI]
Main Effect of Condition				$\chi^2(2) = 18.75$	< 0.001(< 0.001)	partial $\eta^2 = 0.11$ [0.05, 1.00]
Human vs. Control	−0.53	0.1	[−0.74, −0.33]	$t(220) = -5.19$	< 0.001	$d = -0.63$ [−0.98, −0.28]
Human vs. AI	−0.12	0.1	[−0.32, 0.07]	$t(210) = -1.26$	0.208	$d = -0.19$ [−0.49, 0.11]
AI vs. Control	−0.41	0.11	[−0.64, −0.19]	$t(270) = -3.59$	< 0.001	$d = 0.44$ [−0.13, 1.01]

Note. χ^2 = likelihood-ratio test comparing models with vs. without Condition. Model R^2(marginal) = 0.228, R^2(conditional) = 0.286. Adjusted post-study means (SE): Control = 2.37 (0.08), AI = 1.96 (0.08), Human = 1.84 (0.06). p-values in parentheses are Bonferroni-adjusted across nine exploratory outcomes (α = 0.05).

Table A6
Effects of Condition on Post-Study Social Support.

	b	SE	95% CI	$\chi^2(df)$ / $t(df)$	p	partial η^2 / Cohen's d [95% CI]
Main Effect of Condition				$\chi^2(2) = 3.51$	0.173(0.999)	partial $\eta^2 = 0.02$ [0.00, 1.00]
Human vs. Control	0.19	0.11	[−0.03, 0.40]	$t(219) = 1.69$	0.092	$d = 0.21$ [−0.10, 0.52]
Human vs. AI	0.14	0.11	[−0.07, 0.35]	$t(209) = 1.35$	0.179	$d = 0.06$ [−0.29, 0.42]
AI vs. Control	0.04	0.12	[−0.20, 0.28]	$t(271) = 0.36$	0.721	$d = 0.15$ [−0.43, 0.73]

Note. χ^2 = likelihood-ratio test comparing models with vs. without Condition. Model R^2(marginal) = 0.013, R^2(conditional) = 0.122. Adjusted post-study means (*SE*): Control = 2.97 (0.09), AI = 3.01 (0.08), Human = 3.16 (0.07). p-values in parentheses are Bonferroni-adjusted across nine exploratory outcomes (α = 0.05).

Table A7
Effects of Condition on Daily Feeling Heard and Interpersonal Closeness (AI vs. Human Conditions Only).

	b	SE	95% CI	$t(df)$	p	Cohen's d [95% CI]
Feeling Heard	0.16	0.08	[0.00, 0.31]	1.99(161)	0.047(0.442)	0.36 [0.00, 0.72]
Interpersonal Closeness	0.04	0.15	[−0.25, 0.33]	0.28(163)	0.779(0.999)	0.04 [−0.27, 0.36]

Note. $\chi^2(1) = 3.96$, $p = .047$ for Feeling Heard; $\chi^2(1) = 0.08$, $p = .78$ for Interpersonal Closeness (likelihood-ratio tests comparing models with vs. without Condition). Feeling Heard: R^2 (marginal) = 0.012, R^2 (conditional) = 0.590; Interpersonal Closeness R^2 (marginal) = 0, R^2 (conditional) = 0.512. Adjusted means (*SE*): Feeling Heard: AI = 3.84(0.06), Human = 4.00(0.05); Interpersonal Closeness: Human = 2.69(0.10), AI = 2.73(0.11). p-values in parentheses are Bonferroni-adjusted across nine exploratory outcomes (α = 0.05).

Table A8
Effects of Condition on Number of New Friends Made.

	b	SE	95% CI	$t(df)$	p	Cohen's d [95% CI]
Human vs. Control	0.83	0.47	[−0.10, 1.76]	1.75(227)	0.081	0.15 [−0.18, 0.49]
Human vs. AI	0.36	0.45	[−0.53, 1.24]	0.79(215)	0.431	0.11 [−0.17, 0.40]
AI vs. Control	0.47	0.53	[−0.57, 1.52]	0.89(272)	0.374	−0.04 [−0.58, 0.51]

Note. $\chi^2(2) = 3.13$, $p = .209$ (likelihood-ratio test comparing models with vs. without Condition). Table presents post – pre differences in the number of new friends reported (i.e., number at post-study minus number at pre-study).

Appendix B. Reliability and convergent validity of GPT-based content analysis

All reliability analyses were conducted using a balanced subset of 100 daily conversations (50 from the AI condition and 50 from the human condition). Engagement and empathy were rated on 5-point Likert scales. In the human condition, both conversation partners were human; in the AI condition, ratings focused on the human participant's messages. Reliability estimates therefore reflect the consistency of coding for human-generated messages across the two conditions. Across both engagement and empathy, agreement between GPT ratings and the human-coder average was comparable to or higher than agreement between the two human coders.

Table B1
Inter-Rater Reliability and Convergent Validity for Empathy Ratings.

	Weighted κ	ICC	Pearson r
Human–Human coders	0.25	0.25	0.36
Human-average vs. GPT	0.49	0.51	0.52

Note. Human–human reliability reflects agreement between two independent human coders. GPT reliability reflects agreement between GPT ratings and the average of the two human coders. Difference in Pearson correlations was significant (bootstrap $p < .001$).

Table B2
Inter-Rater Reliability and Convergent Validity for Engagement Ratings.

	Weighted κ	ICC	Pearson r
Human–Human coders	0.26	0.26	0.55
Human-average vs. GPT	0.67	0.63	0.66

Note. Human–human reliability reflects agreement between two independent human coders. GPT reliability reflects agreement between GPT ratings and the average of the two human coders. Difference in Pearson correlations was significant (bootstrap $p = .026$).

Appendix C. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jesp.2026.104911.

References

Adewale, M. D., & Muhammad, U. I. (2025). From virtual companions to forbidden attractions: The seductive rise of artificial intelligence love, loneliness, and intimacy—A systematic review. *Journal of Technology in Behavioral Science*, 1–18. https://doi.org/10.1007/s41347-025-00549-4

Arnett, J. J. (2000). Emerging adulthood: A theory of development from the late teens through the twenties. *American Psychologist, 55*(5), 469. https://doi.org/10.1037//0003-066X.55.5.469

Aron, A., Aron, E. N., & Smollan, D. (1992). Inclusion of other in the self scale and the structure of interpersonal closeness. *Journal of Personality and Social Psychology, 63*(4), 596–612. https://doi.org/10.1037/0022-3514.63.4.596

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Brandtzaeg, P. B., Skjuve, M., & Følstad, A. (2022). My AI friend: How users of a social chatbot understand their human–AI friendship. *Human Communication Research, 48*(3), 404–429. https://doi.org/10.1093/hcr/hqac008

Buote, V. M., Pancer, S. M., Pratt, M. W., Adams, G., Birnie-Lefcovitch, S., Polivy, J., & Wintre, M. G. (2007). The importance of friends: Friendship and adjustment among 1st-year university students. *Journal of Adolescent Research, 22*(6), 665–689. https://doi.org/10.1177/0743558407306344

Collins, N. L., & Miller, L. C. (1994). Self-disclosure and liking: A meta-analytic review. *Psychological Bulletin, 116*(3), 457. https://doi.org/10.1037/0033-2909.116.3.457

De Freitas, J., Oğuz-Uğuralp, Z., Uğuralp, A. K., & Puntoni, S. (2025). AI companions reduce loneliness. *Journal of Consumer Research*. , Article ucaf040. https://doi.org/10.1093/jcr/ucaf040

Deci, E. L., La Guardia, J. G., Moller, A. C., Scheiner, M. J., & Ryan, R. M. (2006). On the benefits of giving as well as receiving autonomy support: Mutuality in close friendships. *Personality and Social Psychology Bulletin, 32*(3), 313–327. https://doi.org/10.1177/0146167205282

Dembart, L. (1977). Experts argue whether computers could reason, and if they should. *New York Times, 1*, 34.

Diener, E., Wirtz, D., Tov, W., Kim-Prieto, C., Choi, D., Oishi, S., & Biswas-Diener, R. (2010). New measures of well-being: Flourishing and positive and negative feelings. *Social Indicators Research, 97*(2), 143–156. https://doi.org/10.1007/s11205-009-9493-y

Drouin, M., Sprecher, S., Nicola, R., & Perkins, T. (2022). Is chatting with a sophisticated chatbot as good as chatting online or FTF with a stranger? *Computers in Human Behavior, 128*, Article 107100. https://doi.org/10.1016/j.chb.2021.107100

Epley, N., Waytz, A., & Cacioppo, J. T. (2007). On seeing human: A three-factor theory of anthropomorphism. *Psychological Review, 114*(4), 864–886. https://psycnet.apa.org/doi/10.1037/0033-295X.114.4.864.

Fang, C. M., Liu, A. R., Danry, V., Lee, E., Chan, S. W., Pataranutaporn, P., … Agarwal, S. (2025). *How AI and human behaviors shape psychosocial effects of chatbot use: A longitudinal randomized controlled study*. arXiv preprint arXiv:2503.17473. https://doi.org/10.48550/arXiv.2503.17473.

Folk, D., Yu, S., & Dunn, E. (2024a). Can chatbots ever provide more social connection than humans?. *Collabra. Psychology, 10*(1), Article 117083. https://doi.org/10.1525/collabra.117083

Folk, D., Dunn, E. (2026). How does turning to AI for companionship predict loneliness and vice versa (in press). Psychological Science.

Gable, S. L., Reis, H. T., Impett, E. A., & Asher, E. R. (2004). What do you do when things go right? The intrapersonal and interpersonal benefits of sharing positive events. *Journal of Personality and Social Psychology, 87*(2), 228. https://doi.org/10.1037/0022-3514.87.2.228

Granovetter, M. S. (1973). The strength of weak ties. *American Journal of Sociology, 78*(6), 1360–1380.

Hawkley, L. C., & Cacioppo, J. T. (2010). Loneliness matters: A theoretical and empirical review of consequences and mechanisms. *Annals of Behavioral Medicine, 40*(2), 218–227. https://doi.org/10.1007/s12160-010-9210-8

Helliwell, J. F., Huang, H., Norton, M., Goff, L., & Wang, S. (2023). *World happiness, trust and social connections in times of crisis. World Happiness Report*.

Ho, A., Hancock, J., & Miner, A. S. (2018). Psychological, relational, and emotional effects of self-disclosure after conversations with a chatbot. *Journal of Communication, 68*(4), 712–733. https://doi.org/10.1093/joc/jqy026

Hoffman, M., Richmond, J., Morrow, J., & Salomone, K. (2002). Investigating “sense of belonging” in first-year college students. *Journal of College Student Retention: Research, Theory & Practice, 4*(3), 227–256. https://doi.org/10.2190/DRYC-CXQ9-JQ8V-HT4V

Inzlicht, M., Cameron, C. D., D'Cruz, J., & Bloom, P. (2024). In praise of empathic AI. *Trends in Cognitive Sciences, 28*(2), 89–91. https://doi.org/10.1016/j.tics.2023.12.003

Laurenceau, J. P., Barrett, L. F., & Pietromonaco, P. R. (1998). Intimacy as an interpersonal process: The importance of self-disclosure, partner disclosure, and perceived partner responsiveness in interpersonal exchanges. *Journal of Personality and Social Psychology, 74*(5), 1238. https://doi.org/10.1037/0022-3514.74.5.1238

Laurenceau, J. P., Barrett, L. F., & Rovine, M. J. (2005). The interpersonal process model of intimacy in marriage: A daily-diary and multilevel modeling approach. *Journal of Family Psychology, 19*(2), 314. https://doi.org/10.1037/0893-3200.19.2.314

Lim, M. H., Eres, R., & Vasan, S. (2020). Understanding loneliness in the twenty-first century: An update on correlates, risk factors, and potential solutions. *Social Psychiatry and Psychiatric Epidemiology, 55*, 793–810. https://doi.org/10.1007/s00127-020-01889-7

Lucas, G. M., Gratch, J., King, A., & Morency, L. P. (2014). It's only a computer: Virtual humans increase willingness to disclose. *Computers in Human Behavior, 37*, 94–100. https://doi.org/10.1016/j.chb.2014.04.043

Maisel, N. C., Gable, S. L., & Strachman, A. M. Y. (2008). Responsive behaviors in good times and in bad. *Personal Relationships, 15*(3), 317–338. https://doi.org/10.1111/j.1475-6811.2008.00201.x

Murthy, V. H. (2023). *Our epidemic of loneliness and isolation: The U.S. surgeon general's advisory on the healing effects of social connection and community*. U.S. Department of Health and Human Services. https://www.hhs.gov/sites/default/files/surgeon-general-social-connection-advisory.pdf.

Nass, C., & Moon, Y. (2000). Machines and mindlessness: Social responses to computers. *Journal of Social Issues, 56*(1), 81–103. https://doi.org/10.1111/0022-4537.00153

Perry, A. (2023). AI will never convey the essence of human empathy. *Nature Human Behaviour, 7*(11), 1808–1809. https://doi.org/10.1038/s41562-023-01675-w

R Core Team. (2021). *R: A language and environment for statistical computing (version 4.3.2)* [Computer software]. Vienna, Austria: R Foundation for Statistical Computing https://www.R-project.org/.

Reis, H. T., Clark, M. S., & Holmes, J. G. (2004). Perceived partner responsiveness as an organizing construct in the study of intimacy and closeness. In *Handbook of closeness and intimacy* (pp. 211–236). Psychology Press.

Reis, H. T., & Shaver, P. (1988). Intimacy as an interpersonal process. In S. Duck (Ed.), *Handbook of personal relationships* (pp. 367–389). Wiley.

Roos, C. A., Postmes, T., & Koudenburg, N. (2023). Feeling heard: Operationalizing a key concept for social relations. *PLoS One, 18*(11), Article e0292865. https://doi.org/10.1371/journal.pone.0292865

Russell, D., Peplau, L. A., & Cutrona, C. E. (1980). The revised UCLA loneliness scale: Concurrent and discriminant validity evidence. *Journal of Personality and Social Psychology, 39*(3), 472–480. https://doi.org/10.1037/0022-3514.39.3.472

Shteynberg, G., Halpern, J., Sadovnik, A., Garthoff, J., Perry, A., Hay, J., … Fairweather, A. (2024). Does it matter if empathic AI has no empathy? *Nature Machine Intelligence, 6*(5), 496–497. https://doi.org/10.1038/s42256-024-00841-7

Skjuve, M., Følstad, A., Fostervold, K. I., & Brandtzaeg, P. B. (2021). My chatbot companion-a study of human-chatbot relationships. *International Journal of Human-Computer Studies, 149*, Article 102601. https://doi.org/10.1016/j.ijhcs.2021.102601

Turkle, S. (2011). *Alone together: Why we expect more from technology and less from each other*. Basic Books.

van Wezel, M. M., Croes, E. A., & Antheunis, M. L. (2021). "I'm Here for you": Can SocialChatbots truly support their users? A literature review. *Chatbot research and design: 4th international workshop, CONVERSATIONS 2020, virtual event, November 23–24, 2020. Revised Selected Papers, 4*, 96–113.

World Health Organization. (2023, November 13). WHO launches Commission on Social Connection. Retrieved from: https://www.who.int/news/item/13-11-2023-who-launches-commission-on-social-connection.

Zaki, J. (2019). *The war for kindness: Building empathy in a fractured world*. Crown.