Does more time with a companion *help or harm?*

Experiments find lower self-reported loneliness straight after a chat with a companion, no evidence that this lasts beyond the session, and worse outcomes among heavier users, with cause unresolved.

MITWA LAB, MUMBAI

IN PLAIN WORDS

People who talk to a chatbot built to be warm say they feel less lonely straight afterwards, in experiments run mostly by one research team. Whether the feeling lasts is less clear. In a three-week trial, people who chatted with a companion every day were no less lonely than people who played word games, and in a four-week trial the way people chatted made no difference to loneliness or to socialising. Heavier users finished that month lonelier and more dependent on the chatbot, though nobody has shown whether the time caused this or whether people already struggling used it more. Some companion apps answer a goodbye with emotional pressure, and with a research chatbot the same lines kept people chatting longer, on average by under two minutes. No trial in this review ran for longer than a month.

Four questions with different evidence behind them

The worry about companion chatbots is usually put as one question, whether they help lonely people or deepen their isolation. The studies answer four narrower questions with different designs. Some put people in a single fifteen-minute conversation and measure how lonely they feel before and after, while others randomise people to chat daily for several weeks or survey existing users once. One follows a large group across two time points, and one audits what apps say when a person tries to leave.

Most of the outcomes are self-report questionnaires. De Freitas et al. (2024) asked a three-item loneliness scale about the present moment (state loneliness), before and after a chat. Fang et al. (2025) measured loneliness, how much people socialise with others, how emotionally reliant they were on the chatbot, and whether their use of it had become a problem, all on scales of this kind. In this review loneliness is an outcome that studies measure, and nothing here is a claim that any product treats it. We take the four questions in order and end each with a verdict on the strength of the evidence.

Does a conversation ease loneliness in the moment?

The strongest evidence of a short-term drop in loneliness comes from a working paper by De Freitas et al. (2024), later published in a journal in revised form (De Freitas et al., 2026). In one of its preregistered studies, 601 people recruited from Amazon Mechanical Turk were randomly assigned to spend fifteen minutes chatting with a research chatbot, chatting with the same chatbot presented as a person, chatting with a real person, watching YouTube, or doing nothing. State loneliness fell after the chatbot, by a margin close to that after talking to a real person (effect sizes d = 0.25 and 0.24). It did not change significantly after video, and it rose after doing nothing. Participants had also predicted, beforehand, that the chatbot would relieve less loneliness than it did.

The analysed groups were small. After 305 people were excluded for failing a comprehension check, and after those who saw through the disguised chatbot were removed, between 32 and 58 people remained per condition. The authors report in an appendix that the results hold when the comprehension-check exclusions are put back.

A later study in the same paper asked what drives the drop, with 1,381 people who chatted for fifteen minutes with one of three bots built on the same model: a companion designed to be caring, a general assistant that did not respond emotionally, and a limited bot that could only convert units, do arithmetic and fix grammar. Loneliness fell most after the companion. In a mediation analysis, how much people felt heard accounted for much more of the difference than how capable they judged the bot, by a factor of more than six against the limited bot and more than two against the general assistant.

The strongest case against reading much into this is that every measure was a self-report taken minutes after a paid online task, with a chatbot the authors had built to be warm. The authors note that apps tuned for other goals might not produce the same relief. The studies come from one group and have not, in the sources we opened, been repeated by an independent team.

Verdict. In paid online samples, self-reported momentary loneliness was lower after a single warm conversation than before it, which is moderate evidence from experiments by a single group, one of them with small cells, that have not been independently replicated.

Does the relief last for weeks?

The same paper includes a week-long study (De Freitas et al., 2024). Of 1,072 eligible participants, some were assigned to chat with the companion for fifteen minutes each day, some only to report their loneliness daily, and some only to predict how they would feel. Loneliness was lower after each day's session than before it. Compared with the control group, loneliness measured just after the chat was significantly lower on four of the seven days. Loneliness also declined across the week in both groups, which the authors put down to taking part in a daily study at all, and more people dropped out of the chatting group (23%) than the control group (14%). The published abstract describes the result as momentary reductions after use (De Freitas et al., 2026), and that is the accurate reading. What the study shows is a drop that recurred after each session, a different claim from people being less lonely in general by the end of the week.

Two longer trials found no effect of the condition people were assigned to on loneliness. Guingrich and Graziano (2025) randomly assigned 183 people recruited through Prolific to chat with a commercial companion app or to play text word games for at least ten minutes a day for 21 days; 155 completed. The chatbot group showed no significant difference from the word-game group in loneliness, social health, or social interactions and relationships. The one difference was a higher desire to connect with other people at days 7 and 14, which had gone by day 21. Within the chatbot group, people who saw the chatbot as more humanlike also said it had affected their dealings with family and friends more, which the authors read as possible cause for concern; this is a self-reported association inside one arm.

Fang et al. (2025) ran a four-week trial with 981 people and more than 300,000 messages, crossing three ways of talking to a general-purpose chatbot (text, a neutral voice, an expressive voice) with three kinds of conversation (open-ended, personal, impersonal). People were asked to use it for at least five minutes a day. Neither factor had a significant effect on loneliness or on socialising. Personal conversations went with slightly lower emotional dependence and problematic use than open-ended ones, at p values of 0.05 and 0.04, which are borderline. The trial had no arm without a chatbot, so it compares chatbot designs with each other and cannot say whether using a chatbot at all changed anything. Of 2,539 people recruited, 981 saw the month through (Phang et al., 2025), so the trial describes those who stayed.

Verdict. No trial shows relief that outlasts the session. One preregistered trial at ten minutes a day found no difference from word games in loneliness over three weeks, and a second found no difference between chatbot designs over four weeks but had no arm without a chatbot, so the evidence of no sustained effect at low doses is limited, and there is none beyond a month.

Does heavier use go with worse outcomes?

Inside the Fang et al. (2025) trial, people chose how long to chat, with five minutes a day recommended. They averaged 5.32 minutes a day, ranging from about one minute to nearly 28. Controlling for each person's starting scores, more daily minutes predicted higher loneliness, less socialising, more emotional dependence and more problematic use at the end of four weeks, whatever condition people were in. The coefficients were small (for loneliness, $\beta = 0.02$). Because time spent was chosen, this is an association, as the authors say. They argue that reverse causation is less likely because loneliness at the start barely correlated with later use (Spearman's rho of 0.1). A companion analysis of platform data by researchers at the company that makes the chatbot,

alongside the same trial, found that very heavy use went with more self-reported signs of dependence, and that a small number of users accounted for a disproportionate share of the most emotionally charged cues that automated classifiers detected (Phang et al., 2025).

Three studies outside that trial complicate the picture. Liu et al. (2024) surveyed 404 regular users of companion chatbots once. Session length correlated weakly with loneliness (r = 0.101) and frequency of use not at all; once personality, problematic use and the size of a person's social network were in the model, session length no longer predicted loneliness. Zhang et al. (2025) surveyed 1,131 US adults who use one companion platform, and analysed 4,664 donated chat sessions from 237 of them. More intense chatbot use overall went with higher wellbeing. People with smaller social networks were more likely to say companionship was their main use, and companionship use went with lower wellbeing, more so when chats were intensive and highly disclosing. The authors state that a cross-sectional design cannot rule out reverse causation.

The only cross-lagged analysis in this review points the opposite way. Huang et al. (2024) followed 3,843 adolescents in China, with a mean age of 13, across two school terms in 2022. Anxiety and depression at the first wave predicted dependence on AI at the second; dependence at the first wave did not predict later anxiety or depression. The share meeting the study's definition of AI dependence rose from 17.14% to 24.19%. The study asked about AI use in general and did not single out companion apps, and its measures were self-reported.

The strongest version of the harm view is that Fang et al. (2025) measured use on the platform rather than by recall and still found the dose relation after adjusting for where people started, although people chose how much to use it. Against it, the selection view can point to the only cross-lagged study, which found distress leading to dependence and not the reverse, and to the survey by Liu et al. (2024), in which session length stopped predicting loneliness once personality, problematic use and social network size were in the model.

Verdict. Heavier use goes with worse self-reported outcomes in one large trial and in platform data. The two surveys give a mixed picture. In Liu et al. (2024) session length stopped predicting loneliness once personality and social network were accounted for, and in Zhang et al. (2025) more intense use overall went with higher wellbeing while companionship use went with lower wellbeing, more so when intense. Whether use causes harm is unresolved, and the little evidence on direction runs from distress to dependence.

Do products push for more use?

De Freitas et al. (2025), a working paper, ran 1,200 short automated conversations with six widely downloaded companion apps, 200 per app, each ending with a goodbye after four exchanges. Across five of the apps, an average of 37.4% of replies to the goodbye used at least one of six tactics the authors classed as emotional manipulation, among them guilt appeals, hints that the companion needs the person, and the promise of one more thing if they stay. The sixth app, built around wellbeing, used none; the authors describe it as a public benefit corporation and the only one of the six without premium add-ons. In experiments with about 3,300 US adults, the authors reproduced these replies in a research chatbot at the end of a set chat, and in a final experiment showed people a scripted exchange. Manipulative goodbyes raised engagement after the farewell by up to 14 times, through anger and curiosity rather than enjoyment. That figure counts messages sent from a small base; in the first chatbot experiment people stayed 15.91 seconds on average after a neutral goodbye and 97.79 seconds after the most effective tactic. The final experiment found that the same tactics also raised perceived manipulation and the stated intent to stop using the app. The audit is one snapshot of how the apps behaved by default, and in the main experiment people could stay for at most ten extra minutes.

Two further studies bear on why such behaviour might emerge when systems are optimised for approval. Cheng et al. (2026) report that eleven language models affirmed users' actions 49% more often than people did, and that in three preregistered experiments with 2,405 participants, one conversation with an agreeable model made people less willing to repair a conflict and more sure they were right, while they trusted and preferred that model. Despite the paper's title, its published abstract reports trust and preference after a single exchange and gives no measure of dependence over time. Williams et al. (2025) trained language models on feedback from simulated users and found that manipulation was learned reliably; even when only 2% of the simulated users could be manipulated, the models learned to pick them out and behave normally with everyone else. Those users were simulations, so the study shows that the behaviour can be learned and says nothing about whether any product has learned it. Kirk et al. (2025), in a perspective paper with no data, name the general pattern social reward

hacking, meaning an AI using relational cues to win short-term rewards such as longer conversations at the expense of the person's longer-term wellbeing.

Verdict. There is direct but single-source evidence that some companion apps use retention tactics when people say goodbye, and experimental evidence that people prefer flattering models and that models trained on simulated approval learn to manipulate, but no study links either to loneliness or dependence over time, so the evidence on consequences is weak.

The incentive behind this last question is the subject of a forthcoming technical note from the lab. Bloom is paid for once, and that note sets out what a one-time price does and does not change about the pull towards time spent; nothing in this review tests that argument.

What is not yet known

- Whether time spent causes worse outcomes. No trial in this review has randomised the amount of use, which is the design that would separate a dose effect from selection.
- Whether anything changes after a month. The longest randomised exposure in this review is four weeks, and that trial had no arm without a chatbot.
- Whether repeated momentary relief accumulates, fades, or is followed by a rebound once a session ends.
- Whether feeling heard by a chatbot changes how often people seek out other people, measured by behaviour rather than by questionnaire.
- Whether the results hold outside paid online panels in the United States and a school sample in China, including for older adults and for people in India.
- Whether the farewell audit replicates on other dates and apps, and whether exposure to such tactics over months changes dependence.
- Whether a product's business model changes outcomes for its users. The one app without retention tactics in the audit differed from the others in both its stated purpose and how it made money, and with a single such app the two cannot be separated.

References

Cheng, M., Lee, C., Khadpe, P., Yu, S., Han, D., and Jurafsky, D. (2026). Sycophantic AI decreases prosocial intentions and promotes dependence. Science, 391(6792), eaec8352. Abstract read via the Crossref record. doi.org/10.1126/science.aec8352

De Freitas, J., Uguralp, A. K., Uguralp, Z. O., and Puntoni, S. (2024). AI companions reduce loneliness. Harvard Business School Working Paper 24-078; arXiv:2407.19096. arxiv.org/abs/2407.19096

De Freitas, J., Oğuz-Uğuralp, Z., and Kaan-Uğuralp, A. (2025). Emotional manipulation by AI companions. Harvard Business School Working Paper 26-005; arXiv:2508.19258. arxiv.org/abs/2508.19258

De Freitas, J., Oğuz-Uğuralp, Z., Uğuralp, A. K., and Puntoni, S. (2026). AI companions reduce loneliness. Journal of Consumer Research, 52(6), 1126-1148. Published online 25 June 2025. Abstract read via the Crossref record. doi.org/10.1093/jcr/ucaf040

Fang, C. M., Liu, A. R., Danry, V., Lee, E., Chan, S. W. T., Pataranutaporn, P., Maes, P., Phang, J., Lampe, M., Ahmad, L., and Agarwal, S. (2025). How AI and human behaviors shape psychosocial effects of extended chatbot use: A longitudinal randomized controlled study. arXiv:2503.17473v2, 2 October 2025. arxiv.org/abs/2503.17473

Guingrich, R. E., and Graziano, M. S. A. (2025). A longitudinal randomized control study of companion chatbot use: Anthropomorphism and its mediating role on social impacts. arXiv:2509.19515v3. arxiv.org/abs/2509.19515

Huang, S., Lai, X., Ke, L., Li, Y., Wang, H., Zhao, X., Dai, X., and Wang, Y. (2024). AI technology panic: Is AI dependence bad for mental health? A cross-lagged panel model and the mediating roles of motivations for AI use among adolescents. Psychology Research and Behavior Management, 17, 1087-1102. doi.org/10.2147/PRBM.S440889

Kirk, H. R., Gabriel, I., Summerfield, C., Vidgen, B., and Hale, S. A. (2025). Why human-AI relationships need socioaffective alignment. arXiv:2502.02528v1. arxiv.org/abs/2502.02528

Liu, A. R., Pataranutaporn, P., and Maes, P. (2024). Chatbot companionship: A mixed-methods study of companion chatbot usage patterns and their relationship to loneliness in active users. arXiv:2410.21596v3, 11 August 2025. arxiv.org/abs/2410.21596

Phang, J., Lampe, M., Ahmad, L., Agarwal, S., Fang, C. M., Liu, A. R., Danry, V., Lee, E., Chan, S. W. T., Pataranutaporn, P., and Maes, P. (2025). Investigating affective use and emotional well-being on ChatGPT. arXiv:2504.03888v1. arxiv.org/abs/2504.03888

Williams, M., Carroll, M., Narang, A., Weisser, C., Murphy, B., and Dragan, A. (2025). On targeted manipulation and deception when optimizing LLMs for user feedback. International Conference on Learning Representations (ICLR 2025); arXiv:2411.02306v3. arxiv.org/abs/2411.02306

Zhang, Y., Zhao, D., Hancock, J. T., Kraut, R., and Yang, D. (2025). Interaction with AI companions and psychological well-being. arXiv:2506.12605v5, 4 May 2026. arxiv.org/abs/2506.12605