Intelligent Human Systems Integration (IHSI), Vol. 200, 2026, 340–350
https://doi.org/10.54941/ahfe1007093

Effects of AI Conversational Agents on Stress Reduction: A Meta-Analysis (2015–2025)

Cindy Von Ahlefeldt[1], Ancuta Margondai[2], Valentina Ezcurra[1], Anamaria Acevedo Diaz[1], Soraya Hani[3], and Mustapha Mouloua[1]

[1]College of Sciences, University of Central Florida, Orlando, FL 32816, USA
[2]College of Engineering, University of Central Florida, Orlando, FL 32816, USA
[3]College of Medicine, University of Central Florida, Orlando, FL 32816, USA

ABSTRACT

Between 2015 and 2025, researchers conducted randomized and quasi-experimental trials to investigate whether AI chatbots can reduce perceived stress in adults. These interventions typically delivered CBT, mindfulness, or positive psychology techniques through mobile or web-based platforms. Stress was most commonly assessed using the Perceived Stress Scale (PSS); few studies included physiological measures such as HRV. A review of 34 studies (29 RCTs, 5 quasi-experiments) across diverse populations, such as students, healthcare workers, older adults, and individuals with chronic illnesses, revealed that most interventions lasted between 1 and 16 weeks. About half of the trials reported significantly greater stress reductions in chatbot users compared to controls, generally with small to moderate effect sizes. The remaining studies showed no significant differences. Several studies also reported meaningful improvements in anxiety, depression, or coping self-efficacy, even when stress effects were minimal. The findings suggest that the efficacy of chatbots for stress reduction is mixed and context-dependent, likely moderated by factors such as population, intervention design, and engagement level. Chatbots are seen as promising, scalable tools for stress management; however, future trials should include standardized outcomes, objective stress markers, and longer follow-up periods to better understand their sustained impact.

Keywords: AI chatbots, Conversational agents, Stress reduction, Perceived stress scale (PSS), Mental health interventions, Digital mental health

INTRODUCTION

Chronic stress is a widespread issue among adults and young adults, contributing to anxiety, depression, and adverse physical health outcomes. Traditional stress-management methods such as counseling, CBT, and mindfulness training are effective but often inaccessible due to cost, stigma, and limited availability (Fulmer et al., 2018). In the past decade, AI conversational agents ("chatbots") have emerged as scalable tools for delivering psychoeducational and therapeutic content through naturalistic dialogue (Fitzpatrick et al., 2017). Early studies (2015–2020) indicated potential benefits for mood and anxiety, prompting further investigation into their impact on perceived stress (Fitzpatrick et al., 2017).

Received September 17, 2025; Revised November 10, 2025; Accepted November 25, 2025; Available online February 1, 2026

Stress outcomes in these trials are primarily measured using the Perceived Stress Scale (Cohen et al., 1983), the DASS stress subscale, or visual analog scale. Notably, no trials from 2015–2025 included physiological measures (e.g., HRV, SUDS), underscoring a reliance on subjective assessment.

The growing literature since 2020 includes RCTs testing chatbot efficacy across populations and platforms. This review synthesizes findings from 2015–2025 to determine whether AI chatbots significantly reduce stress, and under what conditions. Key questions include the magnitude of stress reduction, the role of therapeutic modality and population, and the impact on secondary outcomes such as anxiety, coping, and engagement.

BACKGROUND

Chronic stress impacts mental and physical health, increasing risks of conditions like anxiety, depression, and heart disease. Despite effective treatments like CBT and MBSR, barriers such as cost, access, time, and stigma limit utilization, widening the global treatment gap, especially among young, marginalized, and student populations. Digital mental health tools, notably AI chatbots, offer scalable, interactive support that simulates human conversation, delivers personalized feedback, and facilitates self-guided exercises anytime and anywhere. Early research (2015–2020) suggested chatbots may improve mental health outcomes, with recent studies (2020–2025) exploring their potential to reduce stress via cognitive restructuring, mindfulness, and behavioral techniques. Results are mixed, often relying on self-report measures like the PSS with limited biomarker data, making it hard to assess long-term impact. This meta-analysis reviews randomized and quasi-experimental studies (2015–2025) on AI chatbots' effectiveness in reducing perceived stress in adults, aiming to clarify their efficacy and key moderating factors.

METHODS

A comprehensive literature search was conducted for studies published between January 2015 and October 2025. To ensure broad coverage, both keyword-based database queries and a semantic search engine were utilized. The primary search query was: "Effect of AI conversational agents (chatbots) on stress reduction (measured by PSS, SUDS, or HRV) in adults and young adults, 2015–2025, randomized or quasi-experimental." Searches were performed on PubMed, PsycINFO, and Google Scholar using combinations of terms such as "chatbot," "conversational agent," "stress reduction," "Perceived Stress Scale," "randomized trial," among others. Additionally, reference lists of relevant reviews and meta-analyses were hand-searched to identify any studies that may have been missed.

Two independent reviewers screened titles and abstracts against predefined criteria, resolving discrepancies through discussion. The review focused on studies with adult or young adult participants (≥18 years), including college students, working-age adults, and older adults. Interventions targeting minors were excluded for population consistency.

Eligible interventions used an AI-driven conversational agent (chatbot) as the main component of stress or mental health programs. These

chatbots delivered therapeutic or wellness content, such as cognitive behavioral therapy, mindfulness, or coaching, via text or voice and were at least semi-automated. Deliveries could be standalone apps or embedded in messaging platforms, but not entirely therapist-led.

Studies included if they had a control or comparison group, like wait-list, treatment-as-usual, or active controls (e.g., info websites or other apps). Quasi-experimental designs needed a comparison cohort. Studies must include at least one validated stress measure, commonly PSS, SUDS, HRV, DASS-Stress, or visual ratings. Studies measuring only anxiety or depression without stress were excluded.

Only RCTs or quasi-experimental studies with a comparison group were considered; observational or uncontrolled studies were excluded for rigor. The review covered publications from 2015–2025 to track chatbot evolution from rule-based to large language model systems.

A total of 34 studies (29 RCTs, 5 quasi-experiments) met the criteria for full review. Two meta-analyses from 2024–2025 were also examined. The selection process is summarized in a PRISMA diagram.

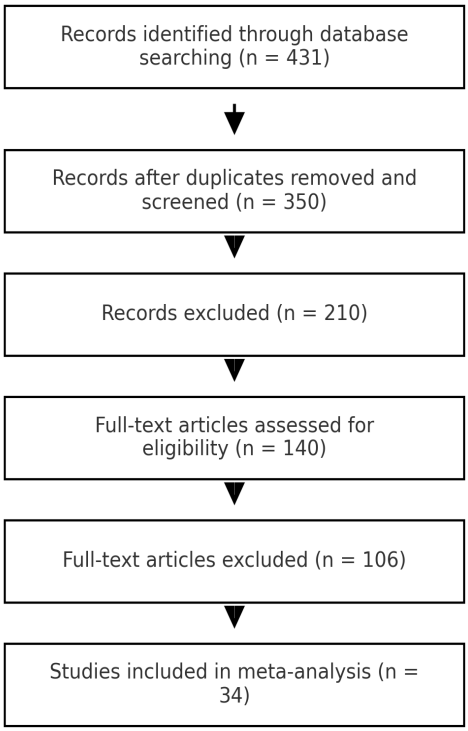

Figure 1: PRISMA flow diagram of study selection (2015–2025).

Data Extraction

For each study, key info on design, population, intervention, outcomes, and engagement was extracted. Design variables included trial type (RCT or quasi-experimental), sample size, attrition, blinding if relevant, and analysis approach. Participant details covered population type (students, employees, patients), age, gender, and inclusion criteria like high stress. Intervention info

listed chatbot name, platform (app, web, messaging, voice), therapeutic style (CBT, mindfulness, etc.), dosage, and automation level. Control conditions noted (waitlist, active, usual care). Primary outcomes involved stress measures (PSS-10, PSS-4, DASS-Stress, visual scales), with baseline, post, follow-up data, differences, p-values, and effect sizes. Secondary outcomes like anxiety, depression, coping, well-being, loneliness were recorded. Engagement metrics and user feedback were also included documented.

Data Synthesis

Due to substantial heterogeneity in populations, interventions, and reporting, a formal meta-analysis was not conducted. Instead, we present a descriptive synthesis of study outcomes. Figure 2 shows individual effect sizes (Hedges' g) for chatbot interventions targeting stress. Several studies (e.g., Ulrich et al., 2023; Ly et al., 2017; Six et al., 2025) demonstrated moderate reductions in stress, while others reported null or minimal effects.

All effects were interpreted using standard benchmarks (Cohen's *d*; 0.2 = small, 0.5 = medium, 0.8 = large). This approach highlights the range and variability of outcomes without imposing a pooled estimate.

RESULTS

Of the 34 studies, 29 were randomized controlled trials and 5 were quasi-experimental. Most RCTs were individually randomized, with sample sizes from fewer than 30 in pilot trials to over 300 in larger studies (median ~100). Participant blinding was often not feasible due to the interactive nature of chatbot interventions; some studies reported blinding of outcome assessors or analysts. The studies involved diverse adult populations. About one-third (n = 11) focused on university students, who tend to report high stress and are familiar with digital tools. Six studies recruited community adults, three involved young adults in stressful situations like post-cancer recovery, and three targeted employees or older workers, including those over 55 with anxiety. Two studies focused on older adults, often using embodied agents for engagement. Two trials recruited women, such as primary care patients or women with war-related anxiety. Other groups included healthcare professionals, chronic illness patients, parents, and those with subclinical anxiety or high stress.

Intervention Types and Platforms

Interventions displayed diverse methods and technology. Thirteen chatbots focused on CBT, addressing cognitive restructuring, behavioral activation, problem solving, and thought reframing. Seven employed mindfulness or MBSR, guiding meditation and breathing. Three centered on positive psychology, like gratitude journaling and optimism. Recent work with large language model chatbots (n = 4; 2023–2025) utilized GPT-3.5 or GPT-4 for natural dialogue, including preoperative anxiety education and self-compassion exercises. Delivery methods included mobile apps (14), web platforms

(10), messaging apps (e.g., Facebook Messenger, WhatsApp, 5), and embodied agents (3D avatars, digital humans, 2). Program durations ranged from single sessions to 8–16 week interventions; some included daily check-ins, others were self-initiated or fixed (e.g., 12 sessions over 6 weeks).

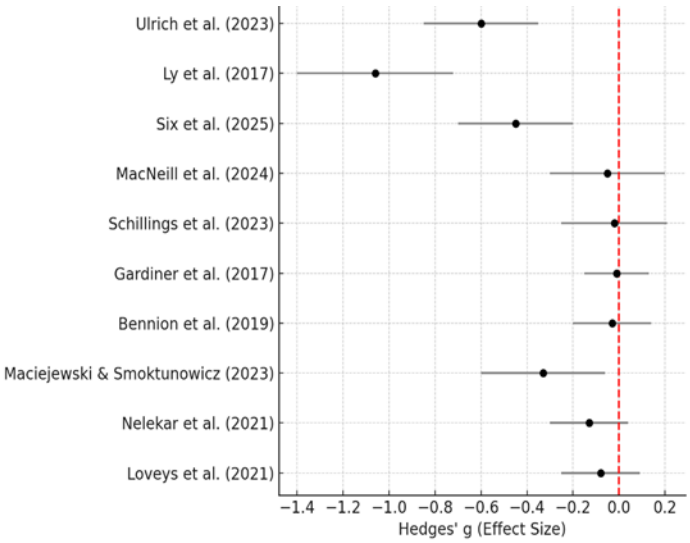

Figure 2: Forest plot of Hedges' g effect sizes for AI chatbot interventions on stress.

All studies included at least one stress measure, mainly the Perceived Stress Scale: 6 used PSS-10, 5 unspecified PSS versions, and 3 used PSS-4. Some used stress subscales from broader measures (e.g., DASS-21) or visual analog scales. None used SUDS or HRV as primary outcomes, relying on subjective self-reports. Many studies included secondary outcomes like anxiety (GAD-7), depression (PHQ-9), coping, resilience, and well-being for broader intervention assessment.

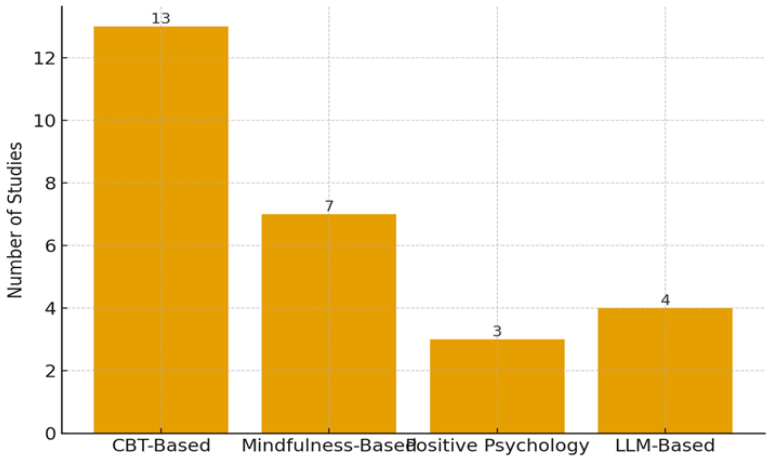

Figure 3: Number of studies by chatbot intervention type (2015–2025).

Primary Outcomes: Stress Reduction Efficacy

Evidence for chatbot-based stress reduction is mixed. About half of controlled trials showed significant stress reduction, while others found no effects. Effect sizes ranged from small to moderate. Several RCTs reported meaningful stress reductions: Ulrich et al. (2023) found a medium effect ($d \approx -0.6$) with a CBT-based chatbot; Ly et al. (2017) saw short-term stress drops with a positive-psychology chatbot, effect sizes from small to large ($d = 0.14–1.06$); Six et al. (2025) observed within-group improvements. Conversely, some trials found no significant effects: MacNeill et al. (2024) saw no differences after four weeks of a CBT chatbot in adults with chronic illness, despite anxiety and depression improvements; Schillings et al. (2023) found no meaningful difference in a three-week mindfulness chatbot trial ($b = -0.018$, $p = .956$); Gardiner et al. (2017) and Bennion et al. (2019) reported minimal differences, possibly due to measurement or placebo effects. Overall, chatbot interventions show potential for short-term stress relief but have inconsistent results across populations and modalities.

Short-Term vs Long-Term Effects

Some interventions produced short-term stress reductions that did not persist over time. For example, Maciejewski & Smoktunowicz (2023) found a modest effect ($d = -0.33$) favoring a self-efficacy chatbot ("Stressbot") at three weeks, but this difference disappeared by one-month follow-up. This pattern of early gains followed by attenuation was observed in several trials, suggesting that sustained engagement or booster content may be required to maintain effects.

Across studies reporting effect sizes ($n \approx 10$), effects ranged from negligible ($d \approx -0.13$) to moderate ($d \approx -0.6$), reflecting considerable variability. Small or nonsignificant effects were reported by Nelekar et al. (2021) and Loveys et al. (2021), while moderate effects favored chatbots in Ulrich et al. (2023) and several pilot studies. Notably, no study found evidence of adverse stress effects, at worst, chatbots showed no difference from control, indicating acceptable safety in this context.

Many trials assessed mental health outcomes like anxiety, depression, and coping. Chatbots often improved anxiety and depression, even when stress effects were nonsignificant. For example, MacNeill et al. (2024) saw reductions in GAD-7 and PHQ-9 among Wysa users despite no change in PSS. Wang et al. (2024) reported moderate effects on depression ($d \approx 0.71$) and loneliness ($d \approx 0.60$) in stressed students, while Young (2025) found GAD-7 reductions up to $d = 1.18$ with an AI mental health app. Reductions in anxiety were observed in Liu et al. (2022) and Sturgill et al. (2020). These suggest chatbots may alleviate stress indirectly by improving mood and worry.

Two recent meta-analyses shed light on chatbot effectiveness. Zhong et al. (2024) analyzed 18 RCTs on anxiety and depression, finding small but significant effects favoring chatbots immediately after intervention (4–8 weeks), which faded by three months—indicating benefits rely on ongoing engagement. Data on stress are limited. Lau et al. (2025) found no significant effect

of AI psychotherapeutic interventions on stress, though benefits for depression and anxiety are consistent. This suggests stress reduction may depend on intervention type, context, or measures adherence.

DISCUSSION

This review shows AI conversational agents as promising stress tools, with effectiveness varying across groups. University students and community groups benefit more, likely due to digital familiarity and lower stress; clinical groups show smaller effects. Older adults may benefit from voice or embodied agents. Structured approaches like CBT or MBSR over weeks tend to be more effective than brief interventions. Effectiveness relies on dosage and engagement; highly engaged users have better outcomes. Null results often stem from short exposure or high dropout, emphasizing the need for sustained interaction.

In some studies, both chatbot and control groups showed similar improvements, suggesting expectancy effects, increased self-monitoring, or the effectiveness of active controls like health information. This points to the need for more robust comparison groups and engaging chatbot designs. Stress measurement often relies on subjective self-reports (e.g., PSS), which may miss short-term or physiological changes. Adding objective measures like HRV and long-term follow-ups can better determine if chatbots provide lasting benefits.

Between 2015 and 2025, the landscape shifted from rule-based bots to large language model (LLM) systems. Newer chatbots may offer more natural, empathetic interactions, but rigorous trials are limited. Future research should compare these systems to earlier models and explore how personalization affects outcomes.

Limitations include small sample sizes, brief follow-ups, and heterogeneity of interventions, which prevent meta-analyses but ensure transparent reporting of findings.

Practically, mental health chatbots offer low-risk, scalable support as early intervention tools or adjuncts. They cannot replace professional treatment for severe or persistent stress. Future research should identify active chatbot components, standardize measures, incorporate physiological data, assess long-term effects, and compare with human-delivered programs. As Karhiy et al. (2024) suggest, chatbot-guided mindfulness may match human coaching, but larger trials are needed.

Theoretical Implications

The pattern of chatbot effects aligns with psychological models. Most interventions, based on cognitive–behavioral frameworks, provide cognitive restructuring and behavioral exercises, explaining the consistent mood improvements (e.g., anxiety and depression). However, the minimal effects on stress suggest chatbots don't target acute stress-coping mechanisms like relaxation or physiological downregulation. They act as synthetic support agents offering appraisal, informational, and emotional support, aligning

with stress-buffering and parasocial theories. Overall, chatbots effectively support cognition and emotion but may need explicit stress-regulation features (e.g., relaxation, mindfulness) to reduce stress levels.

Practical Implications

Chatbots can support mental health care, especially for mood issues. Their easy access through messaging apps helps reach those avoiding traditional therapy due to cost, stigma, or access issues. Surveys show up to 22% of adults have tried chatbots, with nearly half willing to, highlighting their outreach potential. Clinicians may recommend chatbots for psychoeducation, self-monitoring, and early symptom management, saving therapist time for complex cases. Clear safeguards are vital: chatbots lack human judgment and can't detect crises reliably. Backup resources, like hotlines, should be available, and users must know their limits. Developers and health systems should prioritize privacy, safety, and user education to ensure responsible use. Overall, chatbots provide 24/7, stigma-free support that can expand reach but should complement, not replace, professional care.

Framework Advantages

Mental health chatbots' main strength lies in their use of evidence-based frameworks, especially CBT, allowing structured, repeatable interventions like cognitive reframing. They enable personalization and scalability by adjusting language, timing, and support while tracking mood and engagement. Many incorporate strategies like mindfulness and journaling, supporting multi-modal engagement, with about 90% of studies grounded in psychological theory, enhancing validity. Technological advancements, such as NLP and AI, make interactions more responsive and flexible, broadening reach and personalization beyond traditional therapy.

Limitations and Future Directions

These findings should be approached with caution due to limitations such as small, heterogeneous trials often lacking blinding, control groups, or standardized stress measures. Only 17% focused on stress as a primary outcome, reducing confidence in stress results. Self-reported data and diverse chatbot types complicate interpretation. Long-term data are limited. Future research should use larger RCTs with clear stress outcomes, standardized tools, and longer follow-up. Mechanistic and moderator analyses can clarify intervention effects. The potential of generative AI chatbots is promising but needs further study. Ethical, cultural, and safety issues like privacy, crisis response, and inclusivity must be addressed for responsible use.

Addressing the Intervention Finding

Refining chatbot content and deployment is vital for practical use. Given the weak effects on stress, future designs should include explicit stress-management features like mindfulness, breathing exercises, and problem-solving, similar to models like "Mylo." The stronger mood effects imply chatbots

are best as preventive or early-stage tools, especially for those with mild symptoms, with clear escalation to professional care if needed. Engagement depends on empathy and responsiveness; users disengage with impersonal replies. Improving natural language understanding, adding feedback loops, and personalizing interactions can boost retention and effectiveness.

Broader Impact

Conversational AI could broaden mental health support by offering affordable, 24/7, anonymous help, reaching underserved groups like rural residents, shift workers, and veterans. This may reduce stigma and meet rising demand, as nearly half are willing to try a chatbot. However, ethical and equity issues exist: many chatbots lack safety standards, mishandle crises, give misleading feedback, or lack accountability. Privacy, bias, security, and access disparities are risks, worsened by digital literacy, language, and infrastructure gaps. Strong governance and regulation are needed. Developers and policymakers should ensure evidence-based design, monitoring, and crisis protocols. With proper oversight, chatbots could democratize mental health care; without it, they may harm.

CONCLUSION

AI conversational agents from 2015–2025 show potential in reducing perceived stress in adults, though results vary. About half of controlled studies report significant stress reductions with chatbots, while others see no clear benefit. They seem effective at improving mental health outcomes like anxiety and depression, sometimes alongside modest stress relief. Mixed findings may be due to engagement, design, and measurement differences. As chatbots become more advanced, they may become more engaging and personalized, possibly enhancing stress reduction. Currently, they work best as part of broader stress management. Larger, longer-term studies, including meta-analyses, will be needed to confirm their effectiveness. For now, chatbots can help some users under specific conditions but are not a universal solution. Careful design emphasizing evidence-based content and user engagement is key reduction.

REFERENCES

A Luke MacNeill, Doucet, S., & Luke, A. (2024). Effectiveness of a Mental Health Chatbot for People With Chronic Diseases: Randomized Controlled Trial. *JMIR Formative Research*, 8, e50025–e50025. https://doi.org/10.2196/50025

Bennion, M. R., Hardy, G. E., Moore, R. K., Kellett, S., & Millings, A. (2020). Usability, Acceptability, and Effectiveness of Web-Based Conversational Agents to Facilitate Problem Solving in Older Adults: Controlled Study. *Journal of Medical Internet Research*, 22(5), e16794. https://doi.org/10.2196/16794

Danieli, M., Ciulli, T., Mousavi, S. M., & Riccardi, G. (2021). A Participatory Design of Conversational Artificial Intelligence Agents for Mental Healthcare Application (Preprint). *JMIR Formative Research*, 5(12). https://doi.org/10.2196/30053

Danieli, M., Ciulli, T., Mousavi, S. M., Silvestri, G., Barbato, S., Di Natale, L., & Riccardi, G. (2022). Assessing the Impact of Conversational Artificial Intelligence in the Treatment of Stress and Anxiety in Aging Adults: Randomized Controlled Trial. *JMIR Mental Health*, *9*(9), 1–15. https://doi.org/10.2196/38067

Fitzpatrick, K. K., Darcy, A., & Vierhile, M. (2017). Delivering Cognitive Behavior Therapy to Young Adults With Symptoms of Depression and Anxiety Using a Fully Automated Conversational Agent (Woebot): A Randomized Controlled Trial. *JMIR Mental Health*, 4(2). https://doi.org/10.2196/mental.7785

Fulmer, R., Joerin, A., Gentile, B., Lakerink, L., & Rauws, M. (2018). Using Psychological Artificial Intelligence (Tess) to Relieve Symptoms of Depression and Anxiety: Randomized Controlled Trial. *JMIR Mental Health*, *5*(4). https://doi.org/10.2196/mental.9782

Gardiner, P. M., McCue, K. D., Negash, L. M., Cheng, T., White, L. F., Yinusa-Nyahkoon, L., Jack, B. W., & Bickmore, T. W. (2017). Engaging women with an embodied conversational agent to deliver mindfulness and lifestyle recommendations: A feasibility randomized control trial. *Patient Education and Counseling, 100*(9), 1720–1729. https://doi.org/10.1016/j.pec.2017.04.015

Greer, S., Ramo, D., Chang, Y.-J., Fu, M., Moskowitz, J., & Haritatos, J. (2019). Use of the Chatbot "Vivibot" to Deliver Positive Psychology Skills and Promote Well-Being Among Young People After Cancer Treatment: Randomized Controlled Feasibility Trial. *JMIR MHealth and UHealth*, 7(10), e15018. https://doi.org/10.2196/15018

Karhiy, M., Sagar, M., Antoni, M., Loveys, K., & Broadbent, E. (2024). Can a virtual human increase mindfulness and reduce stress? A randomised trial. *Computers in Human Behavior: Artificial Humans*, 2(1), 100069. https://doi.org/10.1016/j.chbah.2024.100069

Lau, Y., How, W., Ang, W. W., Pang, P. C.-I., Wong, S. H., & Chan, K. S. (2025). Artificial Intelligence–Based Psychotherapeutic Intervention on Psychological Outcomes: A Meta-Analysis and Meta-Regression. *Depression and Anxiety*, 2025(1). https://doi.org/10.1155/da/8930012

Liu, H., Peng, H., Song, X., Xu, C., & Zhang, M. (2022). Using AI chatbots to provide self-help depression interventions for university students: A randomized trial of effectiveness. *Internet Interventions*, 27, 100495. https://doi.org/10.1016/j.invent.2022.100495

Loveys, K., Sagar, M., Zhang, X., Fricchione, G. L., & Broadbent, E. (2021). Effects of Emotional Expressiveness of a Female Digital Human on Loneliness, Stress, Perceived Support, and Closeness Across Genders: Randomized Controlled Trial. *Journal of Medical Internet Research*, 23(11), e30624–e30624. https://doi.org/10.2196/30624

Ly, K. H., Ly, A.-M., & Andersson, G. (2017). A fully automated conversational agent for promoting mental well-being: A pilot RCT using mixed methods. *Internet Interventions*, 10, 39–46. https://doi.org/10.1016/j.invent.2017.10.002

Maciejewski, J., & Ewelina Smoktunowicz. (2023). Low-effort internet intervention to reduce students' stress delivered with Meta's Messenger chatbot (Stressbot): A randomized controlled trial. *Internet Interventions*, 33, 100653–100653. https://doi.org/10.1016/j.invent.2023.100653

Nelekar, S., Abdulrahman, A., Gupta, M., & Richards, D. (2021). Effectiveness of embodied conversational agents for managing academic stress at an Indian University (ARU) during COVID-19. *British Journal of Educational Technology*. https://doi.org/10.1111/bjet.13174

Schillings, C., Meißner, E., Erb, B., Bendig, E., Schultchen, D., & Pollatos, O. (2024). Chatbot ELME for Everyday Life Mindfulness Experience: The effects of a chatbot-based intervention on stress and health-related parameters in a stressed sample (Preprint). *JMIR Mental Health*, 11, e50454–e50454. https://doi.org/10.2196/50454

Six, S., Schlesener, E., Hill, V., Babu, S. V., & Byrne, K. (2025). Impact of Conversational and Animation Features of a Mental Health App Virtual Agent on Depressive Symptoms and User Experience Among College Students: Randomized Controlled Trial. *JMIR Mental Health*, 12, e67381–e67381. https://doi.org/10.2196/67381

Sturgill, R., Martinasek, M., Schmidt, T., & Goyal, R. (2021). A Novel Artificial Intelligence-Powered Emotional Intelligence and Mindfulness App (Ajivar) for the College Student Population During the COVID-19 Pandemic: Quantitative Questionnaire Study. *JMIR Formative Research*, *5*(1), e25372. https://doi.org/10.2196/25372

Ulrich, S., Lienhard, N., Hansjörg Künzli, & Kowatsch, T. (2024). MISHA – A Chatbot-delivered Stress Management Coaching for Students: Pilot Randomized Controlled Trial (Preprint). *JMIR Mhealth and Uhealth*, 12, e54945–e54945. https://doi.org/10.2196/54945

Wang, Y., Li, X., Zhang, Q., Yeung, D., & Wu, Y. (2024). Effect of a CBT-Based AI Chatbot on Depression and Loneliness in Chinese University Students: A Randomized Controlled Trial with Financial Stress Moderation (Preprint). *JMIR Mhealth and Uhealth*. https://doi.org/10.2196/63806

Yehuda, C. B., Gilad-Bachrach, R., & Udi, Y. (2024). Improving Engagement and Efficacy of mHealth Micro-Interventions for Stress Coping: An In-The-Wild Study. ArXiv (Cornell University). https://doi.org/10.48550/arxiv.2407.11612

Young, A. (2025). 618. A Randomised Controlled Trial Of An Ai-Enabled Mental Health Intervention For Generalized Anxiety Disorder Symptoms. *The International Journal of Neuropsychopharmacology*, 28(Supplement_2), ii165–ii166. https://doi.org/10.1093/ijnp/pyaf052.328

Zhong, W., Luo, J., & Zhang, H. (2024). The therapeutic effectiveness of artificial intelligence-based chatbots in alleviation of depressive and anxiety symptoms in short-course treatments: A systematic review and meta-analysis. *Journal of Affective Disorders*, 356, 459–469. https://doi.org/10.1016/j.jad.2024.04.057