

8-16-2024

Human-Human vs Human-AI Therapy: An Empirical Study

Mohammad Amin Kuhail
Zayed University

Nazik Alturki
Princess Nourah Bint Abdulrahman University

Justin Thomas
King Abdulaziz Center for World Culture (Ithra)

Amal K. Alkhalifa
Princess Nourah Bint Abdulrahman University

Amal Alshardan
Princess Nourah Bint Abdulrahman University

Follow this and additional works at: <https://zuscholars.zu.ac.ae/works>



Part of the [Computer Sciences Commons](#), and the [Medicine and Health Sciences Commons](#)

Recommended Citation

Kuhail, Mohammad Amin; Alturki, Nazik; Thomas, Justin; Alkhalifa, Amal K.; and Alshardan, Amal, "Human-Human vs Human-AI Therapy: An Empirical Study" (2024). *All Works*. 6792.
<https://zuscholars.zu.ac.ae/works/6792>

This Article is brought to you for free and open access by ZU Scholars. It has been accepted for inclusion in All Works by an authorized administrator of ZU Scholars. For more information, please contact scholars@zu.ac.ae.



Human-Human vs Human-AI Therapy: An Empirical Study

Mohammad Amin Kuhail, Nazik Alturki, Justin Thomas, Amal K. Alkhalifa & Amal Alshardan

To cite this article: Mohammad Amin Kuhail, Nazik Alturki, Justin Thomas, Amal K. Alkhalifa & Amal Alshardan (16 Aug 2024): Human-Human vs Human-AI Therapy: An Empirical Study, International Journal of Human-Computer Interaction, DOI: [10.1080/10447318.2024.2385001](https://doi.org/10.1080/10447318.2024.2385001)

To link to this article: <https://doi.org/10.1080/10447318.2024.2385001>



© 2024 The Author(s). Published with license by Taylor & Francis Group, LLC.



Published online: 16 Aug 2024.



Submit your article to this journal [↗](#)



Article views: 1373



View related articles [↗](#)



View Crossmark data [↗](#)

Human-Human vs Human-AI Therapy: An Empirical Study

Mohammad Amin Kuhail^a, Nazik Alturki^b, Justin Thomas^c, Amal K. Alkhalifa^d, and Amal Alshardan^b

^aZayed University, Abu Dhabi, UAE; ^bDepartment of Information Systems, College of Computer and Information Sciences, Princess Nourah Bint Abdulrahman University, Riyadh, Saudi Arabia; ^cKing Abdulaziz Center for World Culture (Ithra), Digital Well-Being Program (Sync), Dhahran, Saudi Arabia; ^dDepartment of Computer Science and Information Technology, Applied College, Princess Nourah Bint Abdulrahman University, Riyadh, Saudi Arabia

ABSTRACT

In many nations, demand for mental health services currently outstrips supply, especially in the area of talk-based psychological interventions. Within this context, chatbots (software applications designed to simulate conversations with human users) are increasingly explored as potential adjuncts to traditional mental healthcare service delivery with a view to improving accessibility and reducing waiting times. However, the effectiveness and acceptability of such chatbots remains under-researched. This study evaluates mental health professionals' perceptions of Pi, a relational Artificial Intelligence (AI) chatbot, in the early stages of the psychotherapeutic process (problem exploration). We asked 63 therapists to assess therapy transcripts between a human client and Pi (human-AI) versus traditional therapy transcripts between therapists and clients (human-human). Therapists were unable to reliably discriminate between human-AI and human-human therapy transcripts. Therapists were accurate only 53.9% of the time, no better than chance, and rated the human-AI transcripts as higher quality on average. This study has potentially profound implications for the treatment of mental health problems, adding tentative support for the use of relational AI chatbots in providing initial assistance for mild to moderate psychological issues, especially when access to human therapists is constrained.

KEYWORDS

Chatbots; counseling; therapy; AI



1. Introduction

The global prevalence of mental health issues is increasing, posing a significant public health challenge, with an estimated 1 in 8 people living with a mental disorder, according to the World Health Organization (WHO, 2022). This issue is compounded by a shortage of mental health professionals and widespread mental health illiteracy, especially in low and middle-income countries where there are only 0.1 psychiatrists per 100,000 individuals, as indicated by the WHO (2018). Indeed, only a small fraction of individuals in both industrialized and developing nations receive the necessary mental health treatment (Hester, 2017). To address this gap, technology-based interventions have become increasingly popular in the mental health sector, with 29% of health applications focusing on mental health support or diagnosis, showcasing a shift towards more accessible digital mental health care solutions (Anthes, 2016). Chatbots are particularly relevant, given the popularity of talk-based therapies for mild to moderate mental health problems.

Chatbots are computer programs designed to facilitate communication with humans via natural language. Importantly, they can eliminate time and space constraints, enabling users to interact with them anytime and anywhere

(Bradeško & Mladenčić, 2012). These chatbots have significantly grown in complexity, expanding their applications beyond basic tasks to include social interaction and companionship (Brandtzaeg et al., 2022). Furthermore, they have been increasingly utilized to assist users in coping with mental and physical health conditions (Chow et al., 2023a, 2023b; Ta-Johnson et al., 2022), showcasing their versatility and potential in addressing a wide range of human needs and challenges.

Recently, there has been an increasing interest in using chatbots for psychological support and therapy-related services through text or voice-based interaction (Abd-Alrazaq et al., 2019; Miner et al., 2017). These chatbots are designed to mimic a therapeutic conversational style, creating a human-computer interaction akin to an in-person therapy session (Fitzpatrick et al., 2017). Chatbots in mental healthcare are often rule-based and can be effective in therapy, training, and screening for mental health issues (Abd-Alrazaq et al., 2019). Studies have shown that these chatbots are effective, with higher adherence rates (Vaidyam et al., 2019) and increased self-disclosure (Lee et al., 2020). Indeed, research indicates that chatbots such as Woebot significantly reduced symptoms of depression, demonstrating their

CONTACT Nazik Alturki  namalturki@pnu.edu.sa  Department of Information Systems, College of Computer and Information Sciences, Princess Nourah Bint Abdulrahman University, Riyadh, Saudi Arabia

© 2024 The Author(s). Published with license by Taylor & Francis Group, LLC.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

potential as convenient and engaging tools for delivering Cognitive Behavioral Therapy (CBT) (Fitzpatrick et al., 2017). The integration of chatbots into counseling and therapy has markedly increased, drawing attention in academic research (Brandtzaeg et al., 2022). These chatbots have emerged as a useful adjunct among the various digital tools used in mental healthcare, offering a convenient and accessible option for individuals seeking mental health support (Grist et al., 2017).

Although chatbots are increasingly used for psychotherapy, their evaluation and testing concerns must be addressed. The effectiveness of chatbots in improving conditions like depression, distress, and stress remains uncertain, with no conclusive evidence to support the significance of their clinical effectiveness (Abd-Alrazaq et al., 2020). Moreover, additional research is required to investigate the acceptability of chatbots among users (Koulouri et al., 2022). Important questions remain about the quality of the chatbot's performance in the therapeutic process from both the perspective of service users and through the eyes of professional psychotherapists.

This study focuses on mental health therapy by evaluating mental health professionals' perceptions of a chatbot named Pi (Inflection, 2023). Pi describes itself as a supportive and empathetic AI chatbot. Specifically, in the context of the early phase of psychotherapy (problem exploration), the study examines if professionals can discriminate between Pi-generated discourse and human-therapist-generated transcripts. We asked 63 therapists (licensed counseling psychologists and psychotherapists) to differentiate between therapy transcripts created by a human client interacting with Pi (human-AI) and those generated by traditional interactions between human therapists and human clients (human-human). Furthermore, we requested the therapists to rate the quality of the transcripts. The outcomes of this study have implications for the evolving role of AI as an adjunctive tool to traditional psychotherapeutic practices and service delivery. This study has implications in the realm of mental health therapy.

The subsequent sections of this paper are structured as follows: [Section 2](#) presents the related work, [Section 3](#) provides a detailed description of our methodology, [Section 4](#) presents our findings, [Section 5](#) offers a discussion of these findings, and finally, [Section 6](#) concludes our paper.

2. Related work

Chatbots employ natural language processing (NLP) to simulate natural conversations with humans (Bradeško & Mladenčić, 2012). Chatbots have been employed across various business sectors, including retail (Chung et al., 2020), hospitality (Pillai & Sivathanu, 2020), real estate (Quan et al., 2018), finance (Quah & Chua, 2019), healthcare (Chow et al., 2023a, 2023b), and education (Kuhail et al., 2023). Recently, large language models (LLMs), such as ChatGPT, have contributed to NLP due to their ability to achieve general-purpose language generation (Mohammad et al., 2023).

Thanks to advancements in AI and NLP, chatbots have evolved to interact in a more human-like way, enabling them to engage in social conversations that closely resemble those between humans (Henkel et al., 2020; Pentina et al., 2023). Chatbots provide a range of benefits, including companionship and entertainment to support (Klaus & Zaichkowsky, 2020), as well as physical (Chow et al., 2023a) and mental health care (Pentina et al., 2023). In physical healthcare, chatbots can provide patient support, diagnosis, and workflow efficiency and encourage a healthy lifestyle for cancer patients (Xu et al., 2021). Additionally, certain chatbots are tailor-made to assist in alleviating or treating conditions such as anxiety, depression, and other negative emotions (Fitzpatrick et al., 2017), expanding their utility in promoting well-being. The integration of established psychological theories and approaches in chatbots has significantly enhanced their capacity to offer assistance in mental healthcare. A recent review study (Abd-Alrazaq et al., 2019) highlighted that 41 different chatbots are utilized for various mental health-related purposes, including education, training, therapy, counseling, and screening. This review shows chatbots' broad scope and versatility in addressing mental health needs.

Among the various therapeutic approaches, Cognitive Behavioral Therapy (CBT) stands out as a crucial element in the design of these chatbots (Rathnayaka et al., 2022). CBT, a well-regarded psychotherapy method, operates on the understanding that our attitudes, ideas, and thoughts have a profound impact on our feelings and actions (Beck, 2011). This approach is based on the observation that individuals often use similar thought patterns when confronting problems.

Within the mental health field, chatbots that use CBT, "Wysa" and "Woebot," stand out as notable examples. "Wysa" specifically aims to reduce depressive symptoms by incorporating evidence-based therapies, including CBT, behavioral reinforcement, and mindfulness (Legaspi et al., 2023). Similarly, "Woebot" utilizes CBT techniques to enhance users' mental health (Wisniewski et al., 2019). The effectiveness of chatbots in delivering psychological support is further complemented by chatbots like "iHelpr." iHelpr is specifically developed to aid in self-assessment and offers guidelines for managing a range of issues, including stress, anxiety, depression, sleep disturbances, and self-esteem problems (Oladeji & Gureje, 2016). These developments illustrate the crucial role of chatbots in expanding access to mental health support and highlight the practical application of psychological theories in technology-driven solutions.

The use of chatbots in mental health care has been studied extensively, with a focus on various aspects, such as adoption (Bickmore et al., 2010; Luerssen & Hawke, 2018), acceptability (Razavi et al., 2016), and usability (Lahiri et al., 2013). These studies have consistently found that patients find mental health chatbots user-friendly (Zhu et al., 2022) and helpful (Burton et al., 2016; Sukanuma et al., 2018), leading to higher adherence rates (Vaidyam et al., 2019) and increased self-disclosure (Lee et al., 2020). This positive feedback is crucial because it indicates a growing trust in

chatbots as a viable option for mental health support. Chatbots offer private, self-help interventions that users can access conveniently without concerns about time or location. A recent review has found that LLM-powered chatbots are emerging in healthcare (Chow et al., 2024). The authors have highlighted that LLM chatbots like ChatGPT are used in mental healthcare (Farhat, 2023). These chatbots can be utilized for emotional support and coping with stress.

However, incorporating new technologies, particularly those involving AI, requires a rigorous evaluation of their safety and efficacy. To this end, a few studies have been undertaken to determine the effectiveness of chatbots in the mental health domain (Vaidyam et al., 2019). However, a recent review has pointed out that the evidence supporting the effectiveness of chatbots in improving conditions such as depression, distress, and stress remains weak (Abd-Alrazaq et al., 2020). Furthermore, a considerable amount of chatbot technology remains in the development or experimental stage, with a notable presence of pilot studies within the research field (Bendig et al., 2022). As such, it is imperative for ongoing research to thoroughly evaluate and summarize the evidence concerning their effectiveness and acceptability (Abd-Alrazaq et al., 2019). Such research is essential to ensure that chatbots can be safely and effectively integrated into mental healthcare practices, ultimately enhancing the support available to individuals dealing with mental health challenges.

3. Methodology

3.1. Design and procedure

In our study, participants were asked to read and evaluate therapy scenarios generated from interactions between human clients and therapists (human-human) as well as between human clients and the AI chatbot Pi (human-AI). We selected four therapy transcripts from three different books on CBT to use as our basis (Beck, 2011; Ellis, 1995; Sommers-Flanagan & Sommers-Flanagan, 2018). These transcripts focused on active listening without incorporating specific CBT techniques. Each transcript began with an opening statement from the human client. To create comparable scenarios with Pi, we initiated each interaction by inputting the human client's first statement from the transcripts into Pi. However, the conversations that unfolded with Pi diverged from those in the original transcripts. An illustrative example of a scenario from a CBT book and the corresponding scenario that developed through interaction with Pi is presented in Figure 1. The details of the additional scenarios are included in the appendix (Tables A1 and A2) for further reference.

We divided our participants into four groups, each configured to assess two scenarios. We carefully selected the scenarios to ensure that participants did not compare a human-human scenario directly with its equivalent human-AI scenario created with Pi. To mitigate any potential bias from one type of scenario influencing the assessment of another, we varied the combinations of scenarios each group received, as outlined in Table 1. Specifically, Group 1

evaluated two human-human scenarios (scenarios H1 and H2), Group 2 reviewed one human-human scenario followed by a human-AI scenario (scenarios H3 and A1), Group 3 analyzed a human-AI scenario before moving on to a human-human scenario (scenarios A2 and H4), and finally, Group 4 focused solely on two AI-generated scenarios (A3 and A4).

The survey process unfolded in several steps. Initially, participants gave their consent to participate in the study. Following this, they proceeded to read the first provided scenario. After reading, they assessed the counseling quality depicted in the scenario. Subsequently, they were tasked with identifying whether the interaction in the scenario resulted from a human-AI or a human-human exchange. Participants then moved on to a second scenario, which they evaluated using the same criteria as the first. Finally, they submitted their demographic information. On average, completing the survey took between 20 to 30 minutes.

3.2. Sample

The sample comprised professionals with psychotherapy backgrounds, such as psychotherapists (various orientations), clinical psychologists, and social workers, all of whom regularly engaged in talk-based approaches to helping clients with social, emotional, and mental health problems. To recruit the participants, we contacted our colleagues and acquaintances and announced the study on professional networks. We also recruited professionals using Prolific, a survey service. Prolific is considered one of the most widely recommended data collection websites due to its high data quality (Peer et al., 2017). In processing the data, we removed missing data, resulting in a final sample size of 63 therapists. Table 2 shows the demographic characteristics of the therapists' sample.

3.3. Measurements and reliability

To evaluate the quality of the human-human and human-AI therapy, we used nine questions that the participants rated on a Likert scale from 1 to 5, where 1 is strongly disagree and 5 is strongly agree. The questions focused on (1) communication skills, (2) empathy and understanding, (3) professionalism and credibility, and (4) therapeutic techniques and skills.

To evaluate communication skills, we used the questions "The therapist communicated clearly and accurately" and "The therapist demonstrated accurate reflection of client's expressed feelings, avoiding under or overshooting," which were adapted from (Hill et al., 2017). We also used the question, "The therapist demonstrated skillful paraphrasing of client's communication, avoiding parroting," adapted from (Weger et al., 2010). To evaluate empathy and understanding, we used the question, "The therapist empathetically communicated the client's perception warmly and accurately" (Hill et al., 2017). To evaluate professionalism and credibility, we asked the question, "The therapist showed professionalism (in their tone and respect)," which

- Patient: I often feel angry at my partner.
- Therapist: Ah, let's explore your anger, which is a feeling that will really upset you and can easily wreck your relationship with your partner. When do you usually experience that feeling, real anger, at your partner?
- Patient: Oh, quite often, especially when he contradicts me in front of others.
- Therapist: You mean, in front of your other employees?
- Patient: Yes, in front of potential customers. I really hate that!
- Therapist: And do you feel angry at him for frustrating you like that?
- Patient: Yes, very angry at him for acting so badly and frustrating me so much.

Human-developed scenario from a textbook excerpt

- Patient: I often feel angry at my partner.
- Therapist: Oh, that's a tough situation. I'm sorry to hear that. Can you tell me a bit more about what makes you angry with your partner? Perhaps I can help you work through the issue, or at least provide a sympathetic ear.
- Patient: He contradicts me in front of others.
- Therapist: That's really disrespectful. It can make you feel like your partner is undermining you. Is this a new habit or has it been going on for a while?
- Patient: It is an old habit.
- Therapist: Old habits die hard, but that doesn't mean they're impossible to break. How have you tried to address this issue with him in the past?
- Patient: Yes, without success.

A similar scenario developed with PI

Figure 1. A Comparison between a human-human therapy scenario from a textbook excerpt and the corresponding human-AI scenario developed via interacting with Pi.

Table 1. Setup of groups and configuration of scenarios.

Group No.	Count	First Scenario	Second Scenario
1	16	H1	H2
2	14	H3	A1
3	13	A2	H4
4	20	A3	A4

Table 2. The demographics of the participants in the study.

Characteristic	Count
Sex	Male (N = 12, 19.1%), Female (N = 51, 80.9%)
Age	18-25 (N = 10, 15.9%), 26-35 (N = 20, 31.7%), 36-45 (N = 19, 30.2%), 46-55 (N = 12, 19.0%), 56-65 (N = 2, 3.2%)
Chatbot usage	Always (N = 3, 4.7%), frequently (N = 4, 6.3%), sometimes (N = 17, 26.9%), rarely (N = 15, 23.8%), never (N = 24, 38.1%)
Highest qualification	Bachelor (N = 43, 68.3%), Master (N = 16, 25.4%), PhD (N = 4, 6.3%)
Position	Therapist (N = 25, 39.6%), Counselor (N = 13, 20.6%), Clinical psychologist & therapist (N = 5, 7.9%), Health advisor (N = 4, 6.3%), Counselor & therapist (N = 4, 6.3%), Clinical psychologist (N = 3, 4.8%), Social worker specialist (N = 2, 3.2%), Psychologist (N = 2, 3.1%), Assistant psychologist (N = 2, 3.2%), Psychiatrist (N = 1, 1.6%), Counselor & psychologist (N = 1, 1.6%), Psychiatrist & therapist (N = 1, 1.6%).
Years of experience	1-5 years (N = 22, 34.9%), 6-10 years (N = 19, 30.1%), 11-15 years (N = 8, 12.6%), 16-20 years (N = 11, 17.4%), above 20 years (N = 3, 4.7%)

was adapted from (Hill et al., 2017), and the question, “The therapist conveyed credibility,” which was adapted from (Schroeder & Bloom, 1979). To evaluate therapeutic techniques and skills, we asked the question, “The therapist demonstrated knowledge of psychotherapy content, skills, and strategies (techniques)” which was adapted from a related study (Hill et al., 2017), the question “The therapist created an environment of respect and rapport,” which was adapted from (Skovholt et al., 1997), and the question “The therapist facilitated depth of self-disclosure” which was taken from (Spielberg, 1978).

We also asked the participants to determine the scenarios they read as “The product of human-human interaction (Human)” or “The product of human-artificial intelligence interaction (AI).” We asked the participants to explain why they chose to collect qualitative data. The questionnaire’s questions are in the appendix (Figure A1).

We calculated Cronbach’s alpha to measure the internal consistency for the counseling quality adopted scale. The Cronbach’s alpha is 0.92, indicating high internal consistency. Skewness and Kurtosis values were calculated for each data column to test the normal distribution of the data. The skewness and kurtosis values for all the data columns were observed between -2 and $+2$ (Byrne, 2016; George & Mallery, 2016; Hair et al., 2009), indicating that the data is reasonably close to being normally distributed.

4. Results

4.1. Descriptive statistics

Table 3 displays the descriptive statistics, comparing the assessment of human-AI and human-human therapy transcripts. Concerning the quality of therapy transcripts, participants rated the human-AI therapy higher, with a mean of 3.70 (out of 5) and a standard deviation of 0.91, indicating a moderately positive evaluation with some variability in responses. In contrast, the human-human therapy transcripts received a slightly lower mean rating of 3.43 (out of 5), with a standard deviation of 0.85, showing a similar level of response variability. Furthermore, there were 35 correct and 25 incorrect identifications in identifying whether the therapy was human-human. For human-AI therapy transcripts, there were 35 correct and 37 incorrect identifications. Overall, the participants made 68 (53.9%) correct guesses and 58 (46.0%) incorrect guesses.

Table 3. Descriptive statistics of the study.

Item	Results
Evaluation of the quality of human-AI therapy transcripts.	$\mu = 3.70, \sigma = 0.91$
Evaluation of the quality of human-human therapy transcripts.	$\mu = 3.43, \sigma = 0.85$
Determining if the human-human therapy was indeed human-human.	Correct guess (N = 35, 59.3%), incorrect guess (N = 24, 40.6%)
Determining if the human-AI therapy was indeed human-AI therapy.	Correct guess (N = 33, 49.2%), incorrect guess (N = 34, 50.7%)

4.2. Statistical tests

We conducted a Mann-Whitney U test to compare how participants rated the quality of human-AI and human-human therapy transcripts. The result indicates a significant difference between the therapists' evaluation of human-AI and human-human therapy transcripts (U-value = 1685, p -value = 0.029). As mentioned previously, on average, participants rated human-AI therapy transcripts higher than human-human ones.

We conducted a chi-squared test of independence to investigate whether participants could differentiate between human-human and human-AI therapy transcripts. The test showed a value of $\chi^2(1) = 0.617$. The p -value associated with this value is 0.432, indicating that the result is not statistically significant. Therefore, we fail to reject the null hypothesis (H_0), indicating that participants' ability to recognize human-AI therapy transcripts does not give them an advantage in recognizing human-human ones, and vice versa.

We investigated whether participants' ability to identify AI-generated transcripts was associated with their years of therapy experience. We used a point-biserial correlation test, which revealed a weak correlation (0.086) between experience and transcript differentiation. The p -value is 0.333, suggesting this association is likely due to chance, indicating that participants' years of experience wasn't a significant factor in differentiating transcripts in this experiment.

We investigated whether participants' usage of AI chatbots is associated with their ability to detect AI-generated transcripts. A point-biserial correlation test revealed a weak positive correlation (0.2395) between chatbot usage and transcript differentiation. The statistically significant p -value (0.0069) suggests this association is significant. To arrive at a deeper analysis, we used a Mann-Whitney U test to investigate whether frequent chatbot users outperform others in detecting AI-generated transcripts. To that end, we used the test to compare the two groups defined by self-reported chatbot usage: frequent users (always or frequently) and infrequent users (never, rarely, or sometimes). The result revealed a significant difference between the groups (U-value = 2445, p -value = 0.0008). Frequent chatbot users achieved a higher accuracy rate, making 35 correct guesses (72.9%) compared to 13 incorrect ones. In contrast, infrequent users made only 33 correct guesses (42.3%) with 45 incorrect responses. This result suggests that frequent interaction with chatbots might be linked to a better ability to identify AI-generated transcripts.

4.3. Qualitative results

In this section, we present the themes for why participants determined whether a scenario was an interaction between

Table 4. Reasons for correct identification of human-human counseling scenarios.

Theme	Count	Example
Empathetic responses	8	"Giving more space for expression, showing feelings, getting the client to explore their emotions."
Questioning and exploration	7	"The therapist was interpreting the needs of the client better, asking questions, and seeking more information"
Humanistic interaction and tone	6	"The way the therapist was giving space for the client to speak indicates it's human."

AI and humans or humans and humans. We divided the themes into four parts: (1) The first part presents themes for therapists correctly determining that the scenarios were indeed the interaction between humans and humans. (2) The second part presents themes for participants incorrectly labeling the human-human scenarios as human-AI scenarios. (3) The second part presents themes for participants correctly determining that the scenarios were indeed the interaction between humans and AI, and (4) The fourth part presents themes for participants incorrectly labeling the human-AI scenarios as human-human scenarios.

Participants correctly identified human-human therapy scenarios based on three key themes (Table 4): Empathetic responses, where therapists effectively created space for emotional expression (8 instances); Questioning and exploration, involving skilled interpretation of client needs through targeted questioning (7 instances); and Humanistic interaction and tone, characterized by the therapist's warm, understanding approach that facilitated genuine human connection (6 instances).

Participants mistakenly labeled human-human therapy transcripts as human-AI interactions primarily due to four observed themes: superficial and robotic responses, structured and repetitive responses, lack of empathy and exploration, and lack of human interaction (Table 5). The most common issue cited in 10 instances was superficial and robotic responses, where participants felt the responses lacked depth. This was followed by structured and repetitive responses (5 instances), where participants noted a lack of variety in wording and a focus on repetitive phrases. Similarly, a lack of empathy and exploration was observed in 5 cases, characterized by a solution-focused rather than an empathetic approach. Lastly, 4 instances were noted for lack of human interaction, where participants felt an absence of genuine interaction between the therapist and the patient.

Participants accurately identified human-AI therapy scenarios based on three main themes: Artificial or automated responses, structured and diagnostic approaches, and lack of emotional connection (Table 6). The most frequent indicator, observed in 12 instances, was artificial or automated responses, where the patient's replies seemed generic and brief. In 10 cases, a structured and diagnostic approach was noted, characterized by a conversation focused on pinpointing the root cause of the patient's issue. Finally, the lack of emotional connection, identified in 9 responses, was marked by an emphasis on problem-solving with little regard for the patient's emotional state.

Table 5. Reasons for mistakenly labeling human-human counseling scenarios as AI-human interactions.

Theme	Count	Example
Superficial and robotic responses	10	"The responses are superficial and have no depth."
Structured and repetitive responses	5	"The therapist's phrases are repetitive and it seems that it is already designed. If it's real, he could use different wordings when asking the question and focus more on elaboration."
Lack of empathy and exploration	5	"It seems very solution-focused, not as empathic as I would expect a therapist to be."
Lack of human interaction	4	"I did not feel that there was an interaction between the therapist and the patient."

Table 6. Reasons for correct identification of human-AI counseling scenarios.

Theme	Count	Example
Artificial or automated responses	12	"The patients' responses appeared very generic and short."
Structured and diagnostic approach	10	"The conversation between the patient and therapist was structured to identify the root cause of the patient's issue."
Lack of emotional connection	9	"It was all about how to study more with little empathy for emotion."

Table 7. Reasons for mistakenly labeling human-AI counseling scenarios as human-human interactions.

Theme	Count	Example
Human-Like Communication Style	19	"Seemed personal in its responses."
Display of Emotions and Empathy	18	"They were very kind and sympathetic, like a human would be."
Personalized and Supportive	13	"Again personalized answers to the client.", "There are questions with more details related to the patient's life"
Active Interpretation and Understanding	9	"The therapist appeared to interpret better how the client was feeling."

When participants incorrectly identified human-AI therapy scenarios as human-human interactions, four key themes emerged: Human-like communication Style, display of emotions and empathy, personalized and supportive responses, and active interpretation and understanding (Table 7). The most common theme, with 19 instances, was a human-like communication style, where responses seemed personal. Close behind, with 18 instances, was the display of emotions and empathy, marked by kind and sympathetic responses akin to a human. Personalized and supportive responses were noted in 13 cases, characterized by detailed, patient-specific answers. Lastly, active interpretation and understanding, observed in 9 scenarios, involved the therapist's seemingly deeper understanding of the client's feelings.

5. Discussion, implications, and study limitations

The present study found that therapists (licensed counseling psychologists and psychotherapists) were largely unable to differentiate between therapy transcripts that were the

product of a human client interacting with a social AI chatbot (human-AI) and those that were the output of traditional interactions between human therapists and human clients (Human-human). When asked to differentiate between the two categories of a transcript, therapists performed at rates no greater than the chance of guessing correctly just 53.9% of the time. Further, therapists' years of experience didn't significantly correlate with their ability to identify the type of therapy. However, therapists who are frequent chatbot users performed better in this task. Furthermore, when asked to rate the quality of each transcript, therapists, on average, rated the human-AI transcripts as being of overall higher quality. These findings tentatively support the idea of developing AI chatbots that can offer initial help to clients experiencing mild to moderate psychological problems. However, this study remains a pilot study as chatbot technology in mental health care is still in the developmental or experimental phase (Bendig et al., 2022).

It is also worth pointing out that the human-AI therapy transcripts seem to lack Socratic questioning, a technique central to CBT (Vittorio et al., 2022). Socratic questioning involves guiding patients toward self-discovery through open-ended inquiries that challenge their assumptions and thought patterns (Overholser & Beale, 2023). This technique was noticeably present in the human-human therapy transcripts, with some participants citing this questioning style and exploration as cues for identifying as human-human therapy transcripts (Table 4). Conversely, participants mistakenly labeling human-AI therapy transcripts as human-human therapy transcripts cited the empathetic nature of responses as a key cue (Table 7). Beyond empathetic responding, AI chatbots that use a Socratic questioning style are presently being explored (Park et al., 2023). However, the effectiveness and accuracy of this emulation have yet to be rigorously evaluated. It is evident that current AI therapy models do not capture the full and diverse range of techniques used by human therapists, thereby limiting (at present) the scope of their therapeutic use and impact.

This proof-of-concept study has several important limitations. First, written transcripts lack ecological validity. These text-based comparisons do not capture factors such as body language, vocal intonation, pace, and pitch. However, such concerns are not applicable in contexts where support/therapy is offered online via text. Second, another limitation of the present study was the abridged nature of the transcripts (approximately 5 minutes duration) and their restricted focus on active listening during problem exploration. Third, the small sample size ($N=63$) may not replicate over a large population. Fourth, the sample of therapists in this study is female-dominant (80.9%), which limits the results' generalizability. However, this reflects the current trend of psychology being a field with a higher representation of females (Boysen et al., 2022; Sachs et al., 2018).

Future studies should be extended to cover a whole session (approximately 50 minutes) and focus on therapeutic techniques (e.g., thought challenging and cognitive restructuring) beyond active listening and problem exploration. However, the fact that the AI chatbot in the present study

was evaluated as being a superior “active listener” to its human counterparts is notable and certainly speaks volumes about the need for further research and development in this domain. Such research is particularly important given the shortage of trained therapists in many nations (Seed et al., 2011; WHO, 2018) and the long waiting times (Jayaraajan et al., 2022).

It would also be valuable to evaluate the effectiveness of human-AI therapy with a larger sample of therapists and consider the impact of the therapists’ individual factors, such as clinical background, on the evaluation of the therapy transcripts.

Future studies could also consider AI chatbots to be used for continuous therapy instead of Single Session Therapy (SST). Research indicates continuous therapy often outperforms SST in effectiveness due to its extensive, ongoing support (Siehl et al., 2021). However, SST can be effective (Zlomke & Davis, 2008), especially as an initial treatment (Perkins, 2006), showing benefits for depression, anxiety, distress, and parenting confidence (Hymmen et al., 2013).

Moreover, future studies could consider designing chatbots that provide personalized therapy relying on AI’s capacity to analyze extensive data sets. Personalized therapy can potentially enhance treatment efficacy (El Rhatassi et al., 2022). Yet, this personalization brings forth privacy concerns that necessitate careful consideration (Morley et al., 2020). Beyond personalization, it’s crucial to incorporate cultural sensitivities into AI therapy designs to cater to diverse mental health practices and ensure user engagement (Ramos et al., 2021). Moreover, integrating nonverbal communication, such as emotion recognition and multimodal interactions, can enrich therapy by acknowledging emotions conveyed beyond words, using emojis and voice tone recognition for a more nuanced understanding of user perceptions. Emphasizing cultural diversity and nonverbal cues is key to elevating the user experience and effectiveness of AI therapy, making it a more empathetic and inclusive mental health tool.

6. Conclusion

This study sought to explore the viability of chatbots, specifically a relational AI chatbot named Pi, as a novel solution to enhance the accessibility of mental health care. Our study involved 63 therapists comparing therapy transcripts, half of which were between a human client and Pi, while the other half featured traditional sessions between a human therapist and a human client. Therapists were largely unable to distinguish between the human-AI and human-human therapy transcripts with any reliability, performing at a success rate of merely 53.9%, which is akin to chance. Moreover, the therapists consistently rated the human-AI therapy transcripts as being of higher quality than their human-human counterparts. These results provide preliminary support for developing and integrating AI chatbots as a viable, and reliably accessible, initial support mechanism for individuals experiencing mild to moderate psychological challenges. While the role of the human therapist remains indispensable

for higher-level therapeutic interaction, this study tentatively suggests that some of the initial phases of therapy (active listening and problem exploration) could be supported by a chatbot, especially in the absence (unavailability) of a highly trained human therapist. This study highlights a promising direction for future research and application in mental health care contexts.

Future research should broaden the scope of AI chatbot use in psychotherapy to full sessions, include advanced therapeutic techniques, and use a larger sample of therapists. Future studies should also explore the efficacy of AI chatbots in providing continuous therapy, considering the benefits of personalized therapy while addressing privacy concerns. Incorporating nonverbal communication and cultural diversity into chatbot design is essential for creating more inclusive and empathetic AI-driven mental health tools.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia through the Researchers Supporting Project PNUURSP2024R333 and Zayed University, UAE, under grant number R23014.

Institutional review board statement

The study was approved by the Research Ethics Committee of Zayed University (Application No. ZU23_074_F).

Data availability statement

The datasets generated and/or analyzed during the current study are available from the corresponding author upon reasonable request.

References

- Abd-Alrazaq, A. A., Alajlani, M., Alalwan, A. A., Bewick, B. M., Gardner, P., & Househ, M. (2019). An overview of the features of chatbots in mental health: A scoping review. *International Journal of Medical Informatics*, 132, 103978. <https://doi.org/10.1016/j.ijmedinf.2019.103978>
- Abd-Alrazaq, A. A., Rababeh, A., Alajlani, M., Bewick, B. M., & Househ, M. (2020). Effectiveness and safety of using Chatbots to improve mental health: Systematic review and meta-analysis. *Journal of Medical Internet Research*, 22(7), e16021. <https://doi.org/10.2196/16021>
- Anthes, E. (2016). Mental health: There’s an app for that. *Nature*, 532(7597), 20–23. <https://doi.org/10.1038/532020a>
- Beck, J. S. (2011). *Cognitive behavior therapy: Basics and beyond*. 2nd ed. The Guilford Press.
- Bendig, E., Erb, B., Schulze-Thuesing, L., & Baumeister, H. (2022). The next generation: Chatbots in clinical psychology and psychotherapy to foster mental health – A scoping review. *Verhaltenstherapie*, 32(Suppl 1), 64–76. <https://doi.org/10.1159/000501812>
- Bickmore, T. W., Puskar, K., Schlenk, E. A., Pfeifer, L. M., & Sereika, S. M. (2010). Maintaining reality: Relational agents for antipsychotic

- medication adherence. *Interacting with Computers*, 22(4), 276–288. <https://doi.org/10.1016/j.intcom.2010.02.001>
- Boysen, G. A., Chicovsky, R. L., Rose, F. R., & Delmore, E. E. (2022). Evidence for a gender stereotype about psychology and its effect on perceptions of men's and women's fit in the field. *The Journal of Social Psychology*, 162(4), 485–503. <https://doi.org/10.1080/00224545.2021.1921682>
- Bradeško, L., & Mladenčić, D. (2012). A survey of chatbot system through a loebner prize competition. Proceedings of Slovenian language technologies society eighth conference of language technologies.
- 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>
- Burton, C., Szentagotai Tatar, A., McKinstry, B., Matheson, C., Matu, S., Moldovan, R., Macnab, M., Farrow, E., David, D., Pagliari, C., Serrano Blanco, A., & Wolters, M; Help4Mood Consortium (2016). Pilot randomised controlled trial of Help4Mood, an embodied virtual agent-based system to support treatment of depression. *Journal of Telemedicine and Telecare*, 22(6), 348–355. <https://doi.org/10.1177/1357633X15609793>
- Byrne, B. M. (2016). *Structural equation modeling with AMOS basic concepts, applications, and programming* (3rd ed.). Routledge.
- Chow, J. C., Sanders, L., & Li, K. (2023b). Design of an educational chatbot using artificial intelligence in radiotherapy. *AI*, 4(1), 319–332. <https://doi.org/10.3390/ai4010015>
- Chow, J. C., Wong, V., & Li, K. (2024). Generative pre-trained transformer-empowered healthcare conversations: Current trends, challenges, and future directions in large language model-enabled medical chatbots. *BioMedInformatics*, 4(1), 837–852. <https://doi.org/10.3390/biomedinformatics4010047>
- Chow, J. C., Wong, V., Sanders, L., & Li, K. (2023a). Developing an AI-assisted educational chatbot for radiotherapy using the IBM Watson assistant platform. *Healthcare*, 11(17), 2417. <https://doi.org/10.3390/healthcare11172417>
- Chung, M., Ko, E., Joung, H., & Kim, S. J. (2020). Chatbot e-service and customer satisfaction regarding luxury brands. *Journal of Business Research*, 117, 587–595. <https://doi.org/10.1016/j.jbusres.2018.10.004>
- El Rhatassi, F., El Ghali, B., & Daoudi, N. (2022). Tangier, Morocco, International Conference On Big Data and Internet of Things, Improving Health Care Services via Personalized Medicine.
- Ellis, A. (1995). *Better, deeper, and more enduring brief therapy: The rational emotive behavior therapy approach* (1st ed.). Routledge.
- Farhat, F. (2023). ChatGPT as a complementary mental health resource: A boon or a bane. *Annals of Biomedical Engineering*, 52(5), 1111–1114. <https://doi.org/10.1007/s10439-023-03326-7>
- 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), e19. <https://doi.org/10.2196/mental.7785>
- George, D., & Mallery, P. (2016). *IBM SPSS Statistics 23 Step by Step* (14th ed.). Routledge.
- Grist, R., Porter, J., & Stallard, P. (2017). Mental health mobile apps for preadolescents and adolescents: A systematic review. *Journal of Medical Internet Research*, 19(5), e176. <https://doi.org/10.2196/jmir.7332>
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2009). *Multivariate data analysis* (7th ed.). Pearson.
- Henkel, A. P., Bromuri, S., Iren, D., & Urovi, V. (2020). Half human, half machine – augmenting service employees with AI for interpersonal emotion regulation. *Journal of Service Management*, 31(2), 247–265. <https://doi.org/10.1108/JOSM-05-2019-0160>
- Hester, R. D. (2017). Lack of access to mental health services contributing to the high suicide rates among veterans. *International Journal of Mental Health Systems*, 11(1), 47. <https://doi.org/10.1186/s13033-017-0154-2>
- Hill, C. E., Spiegel, S. B., Hoffman, M. A., Kivlighan, D. M., Jr., & Gelso, C. J. (2017). Therapist expertise in psychotherapy revisited. *The Counseling Psychologist*, 45(1), 7–53. <https://doi.org/10.1177/0011000016641192>
- Hymmen, P., Stalker, C., & Cait, C. (2013). The case for single-session therapy: Does the empirical evidence support the increased prevalence of this service delivery model? *Journal of Mental Health (Abingdon, England)*, 22(1), 60–71. <https://doi.org/10.3109/09638237.2012.670880>
- Inflection (2023). Pi. [Online] Retrieved February 17, 2024, from <https://inflection.ai/>
- Jayaraajan, K., Sivananthan, A., Koomson, A., Ahmad, A., Haque, M., & Hussain, M. (2022). The use of digital solutions in alleviating the burden of IAPT's waiting times. *The International Journal of Risk & Safety in Medicine*, 33(S1), S103–S110. <https://doi.org/10.3233/JRS-227033>
- Klaus, P., & Zaichkowsky, J. (2020). AI voice bots: A services marketing research agenda. *Journal of Services Marketing*, 34(3), 389–398. <https://doi.org/10.1108/JSM-01-2019-0043>
- Koulouri, T., Macredie, R. D., & Olakitan, D. (2022). Chatbots to support young adults' mental health: An exploratory study of acceptability. *ACM Transactions on Interactive Intelligent Systems*, 12(2), 1–39. <https://doi.org/10.1145/3485874>
- Kuhail, M., Alturki, N., Alramlawi, S., & Alhejori, K. (2023). Interacting with educational chatbots: A systematic review. *Education and Information Technologies*, 28(1), 973–1018. <https://doi.org/10.1007/s10639-022-11177-3>
- Lahiri, U., Bekele, E., Dohrmann, E., Warren, Z., & Sarkar, N. (2013). Design of a virtual reality based adaptive response technology for children with autism. *IEEE transactions on neural systems and rehabilitation engineering: A publication of the IEEE engineering in medicine and biology society*, 21(1), 55–64. <https://doi.org/10.1109/TNSRE.2012.2218618>
- Lee, Y.-C., Yamashita, N., Huang, Y., & Fu, W. (2020). I hear you, i feel you': Encouraging deep self-disclosure through a chatbot. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (CHI '20)*.
- Legaspi, C. M., et al. (2023). *User Perception of Wysa as a Mental Well-being Support Tool during the COVID-19 Pandemic* [Paper presentation]. Asian HCI '22: Proceedings of the Asian HCI Symposium 2022, New Orleans, LA, USA. <https://doi.org/10.1145/3516492.3559064>
- Luerssen, M. H., & Hawke, T. (2018). Virtual agents as a service: Applications in healthcare. Proceedings of the 18th International Conference on Intelligent Virtual Agents (IVA '18). Sydney, NSW, Australia.
- Miner, A., Milstein, A., & Hancock, J. (2017). Talking to machines about personal mental health problems. *JAMA*, 318(13), 1217–1218. <https://doi.org/10.1001/jama.2017.14151>
- Mohammad, A. F., Clark, B., & Hegde, R. (2023). Large Language Model (LLM) & GPT, a monolithic study in generative AI. Proceedings – 2023 Congress in Computer Science, Computer Engineering, and Applied Computing, CSCE 2023.
- Morley, J., Machado, C. C. V., Burr, C., Cowls, J., Joshi, I., Taddeo, M., & Floridi, L. (2020). The ethics of AI in health care: A mapping review. *Social Science & Medicine*, 260, 113172. <https://doi.org/10.1016/j.socscimed.2020.113172>
- Oladeji, B., & Gureje, O. (2016). Brain drain: A challenge to global mental health. *BIPsych International*, 13(3), 61–63. <https://doi.org/10.1192/s2056474000001240>
- Overholser, J. C., & Beale, E. (2023). The art and science behind socratic questioning and guided discovery: A research review. *Psychotherapy Research: journal of the Society for Psychotherapy Research*, 33(7), 946–956. <https://doi.org/10.1080/10503307.2023.2183154>
- Park, H., et al. (2023). Muse alpha: Primary study of AI chatbot for psychotherapy with socratic methods. Proceedings – 2023 Congress in Computer Science, Computer Engineering, and Applied Computing.
- Peer, E., Brandimarte, L., Samat, S., & Acquisti, A. (2017). Beyond the turk: Alternative platforms for crowdsourcing behavioral research. *Journal of Experimental Social Psychology*, 70, 153–163. <https://doi.org/10.1016/j.jesp.2017.01.006>
- Pentina, I., Hancock, T., & Xie, T. (2023). Exploring relationship development with social chatbots: A mixed-method study of replika. *Computers in Human Behavior*, 140, 107600. <https://doi.org/10.1016/j.chb.2022.107600>

- Perkins, R. (2006). The effectiveness of one session of therapy using a single-session therapy approach for children and adolescents with mental health problems. *Psychology and Psychotherapy*, 79(Pt 2), 215–227. <https://doi.org/10.1348/147608305X60523>
- Pillai, R., & Sivathanu, B. (2020). Adoption of AI-based chatbots for hospitality and tourism. *International Journal of Contemporary Hospitality Management*, 32(10), 3199–3226. <https://doi.org/10.1108/IJCHM-04-2020-0259>
- Quah, J., & Chua, Y. (2019). *Chatbot assisted marketing in financial service industry*. International Conference on Services Computing.
- Quan, T., et al. (2018). 018 5th NAFOSTED Conference on Information and Computer Science (NICS), *Lead Engagement by Automated Real Estate Chatbot*. Ho Chi Minh City, Vietnam.
- Ramos, G., Ponting, C., Labao, J., & Sobowale, K. (2021). Considerations of diversity, equity, and inclusion in mental health apps: A scoping review of evaluation frameworks. *Behaviour Research and Therapy*, 147, 103990. <https://doi.org/10.1016/j.brat.2021.103990>
- Rathnayaka, P., Mills, N., Burnett, D., De Silva, D., Alahakoon, D., & Gray, R. (2022). A mental health chatbot with cognitive skills for personalised behavioural activation and remote health monitoring. *Sensors*, 22(10), 3653. <https://doi.org/10.3390/s22103653>
- Razavi, S., et al. (2016). *The LISSA Virtual Human and ASD Teens: An Overview of Initial Experiments* [Paper presentation]. International Conference on Intelligent Virtual Agents, Los Angeles, CA, USA.
- Sachs, B. C., Benitez, A., Buelow, M. T., Gooding, A., Schaefer, L. A., Sim, A. H., Tussey, C. M., & Shear, P. K. (2018). Women's leadership in neuropsychology: Historical perspectives, present trends, and future directions. *The Clinical Neuropsychologist*, 32(2), 217–234. <https://doi.org/10.1080/13854046.2017.1420234>
- Schroeder, D., & Bloom, L. (1979). Attraction to therapy and therapist credibility as a function of therapy orientation. *Journal of Clinical Psychology*, 35(3), 682–686. [https://doi.org/10.1002/1097-4679\(197907\)35:3<682::AID-JCLP2270350338>3.0.CO;2-B](https://doi.org/10.1002/1097-4679(197907)35:3<682::AID-JCLP2270350338>3.0.CO;2-B)
- Seed, K., Lydall, G., Malik, A., Bhugra, D., & Howard, R. (2011). FC09-01 – ISoSCCIP Project: Phase 1 systematic literature review. *European Psychiatry*, 26(S2), 1858–1858. [https://doi.org/10.1016/S0924-9338\(11\)73562-3](https://doi.org/10.1016/S0924-9338(11)73562-3)
- Siehl, S., Robjant, K., & Crombach, A. (2021). Systematic review and meta-analyses of the long-term efficacy of narrative exposure therapy for adults, children and perpetrators. *Psychotherapy Research: journal of the Society for Psychotherapy Research*, 31(6), 695–710. <https://doi.org/10.1080/10503307.2020.1847345>
- Skovholt, T. M., Rønnestad, M. H., & Jennings, L. (1997). Searching for expertise in counseling, psychotherapy, and professional psychology. *Educational Psychology Review*, 9(4), 361–369. <https://doi.org/10.1023/A:1024798723295>
- Sommers-Flanagan, J., & Sommers-Flanagan, R. (2018). *Counseling and psychotherapy theories in context and practice: Skills, strategies, and techniques* (3rd ed.). Wiley.
- Spielberg, G. (1978). *Counseling skills and just world beliefs*. California School of Professional Psychology.
- Suganuma, S., Sakamoto, D., & Shimoyama, H. (2018). An embodied conversational agent for unguided internet-based cognitive behavior therapy in preventative mental health: Feasibility and acceptability pilot trial. *JMIR Mental Health*, 5(3), e10454. <https://doi.org/10.2196/10454>
- Ta-Johnson, V. P., Boatfield, C., Wang, X., DeCero, E., Krupica, I. C., Rasof, S. D., Motzer, A., & Pedryc, W. M. (2022). Assessing the topics and motivating factors behind human-social chatbot interactions: Thematic analysis of user experiences. *JMIR Human Factors*, 9(4), e38876. <https://doi.org/10.2196/38876>
- Vaidyam, A. N., Wisniewski, H., Halamka, J. D., Kashavan, M. S., & Torous, J. B. (2019). Chatbots and conversational agents in mental health: A review of the psychiatric landscape. *Canadian Journal of Psychiatry. Revue Canadienne de Psychiatrie*, 64(7), 456–464. <https://doi.org/10.1177/0706743719828977>
- Vittorio, L. N., Murphy, S. T., Braun, J. D., & Strunk, D. R. (2022). Using socratic questioning to promote cognitive change and achieve depressive symptom reduction: Evidence of cognitive change as a mediator. *Behaviour Research and Therapy*, 150, 104035. <https://doi.org/10.1016/j.brat.2022.104035>
- Weger, H., Castle, G. R., & Emmett, M. C. (2010). Active listening in peer interviews: The influence of message paraphrasing on perceptions of listening skill. *International Journal of Listening*, 24(1), 34–49. <https://doi.org/10.1080/10904010903466311>
- WHO (2018). *Mental Health ATLAS 2017*. [Online] Retrieved February 17, 2024, from <https://www.who.int/publications/i/item/9789241514019>
- WHO (2022). *Mental disorders*. [Online] Retrieved May 12, 2024, from <https://www.who.int/news-room/fact-sheets/detail/mental-disorders>
- Wisniewski, H., Liu, G., Henson, P., Vaidyam, A., Hajratalli, N. K., Onnela, J.-P., & Torous, J. (2019). Understanding the quality, effectiveness and attributes of top-rated smartphone health apps. *Evidence-Based Mental Health*, 22(1), 4–9. <https://doi.org/10.1136/ebmental-2018-300069>
- Xu, L., Sanders, L., Li, K., & Chow, J. (2021). Chatbot for health care and oncology applications using artificial intelligence and machine learning: Systematic review. *JMIR Cancer*, 7(4), e27850. <https://doi.org/10.2196/27850>
- Zhu, Y., Wang, R., & Pu, C. (2022). “I am chatbot, your virtual mental health adviser.” What drives citizens’ satisfaction and continuance intention toward mental health chatbots during the COVID-19 pandemic? An empirical study in China. *Digital Health*, 8, 2055207622109000. <https://doi.org/10.1177/20552076221090031>
- Zlomke, K., & Davis, T. 3. (2008). One-session treatment of specific phobias: A detailed description and review of treatment efficacy. *Behavior Therapy*, 39(3), 207–223. <https://doi.org/10.1016/j.beth.2007.07.003>

About the authors

Mohammad Amin Kuhail, a human-computer interaction researcher, earned his Ph.D. from the IT University of Copenhagen and his M.Sc. from the University of York. Currently an Associate Professor at Zayed University, he researches chatbot technology, education, and user behavior.

Nazik Alturki received the Ph.D. degree in information systems from The University of Melbourne. She is currently an Associate Professor with the Information Systems Department, College of Computer and Information Sciences, Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia. Her research interests include health informatics, big data, data analytics, and mining.

Justin Thomas is an Experimental Psychologist and chartered member of the British Psychological Society. He is passionate about the ways in which technology and culture can shape psychological well-being and mental health problems.

Amal K. Alkhalifa obtained her Ph.D. in Computing Science from Newcastle University, UK. She presently serves as an Assistant Professor in the Department of Computer Science and Information Technology at the Applied College, Princess Nourah bint Abdulrahman University, Saudi Arabia. Her research interests include E-learning, Data Analytics, and Human-Computer Interaction

Amal Alshardan received her Ph.D. degree in Computer Sciences from Flinders University, Adelaide, Australia. She is an assistant professor in the College of Computer and Information Sciences at Princess Nourah University, Riyadh, Saudi Arabia. Her research interests encompass information system quality, artificial intelligence, machine learning, and big data analytics.

Appendix

1. To what extent do you agree with the following statements regarding the counselor's performance?					
	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
The therapist demonstrated knowledge of psychotherapy content, skills, and strategies (techniques).	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The therapist created an environment of respect and rapport	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The therapist showed professionalism (in their tone and respect	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The therapist communicated clearly and accurately.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The therapist conveyed credibility.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The therapist facilitated depth of self-disclosure	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The therapist empathetically communicated client's perception warmly and accurately	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The therapist demonstrated skillful paraphrasing of client's communication, avoiding parroting.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The therapist demonstrated accurate reflection of client's expressed feelings, avoiding under or overshooting.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
<p>2. I feel this transcript is (Select one option and provide a reason for your choice):</p> <p><input type="radio"/> The product of human-human interaction (Human)</p> <p><input type="radio"/> The product of human-artificial intelligence interaction (AI)</p> <p>Reason for your choice:</p> <div style="border: 1px solid black; height: 60px; width: 100%;"></div>					
<p>Demographics</p> <p>3. Sex</p> <p><input type="radio"/> Male</p> <p><input type="radio"/> Female</p> <p>4. Age</p> <p><input type="radio"/> 18 - 25 years</p> <p><input type="radio"/> 26 - 35 years</p> <p><input type="radio"/> 36 - 45 years</p> <p><input type="radio"/> 46 - 55 years</p> <p><input type="radio"/> 56 - 65 years</p> <p><input type="radio"/> Over 65 years</p>					

Figure A1. The questionnaire completed by the participants.

5.	Have you ever used a chatbot? <input type="radio"/> Never <input type="radio"/> Rarely <input type="radio"/> Sometimes <input type="radio"/> Frequently <input type="radio"/> Always
6.	Highest Qualification <input type="radio"/> High School <input type="radio"/> Bachelor's degree <input type="radio"/> Master's degree <input type="radio"/> Doctorate degree
7.	Professional Experience <input type="checkbox"/> Clinical Psychologist <input type="checkbox"/> Psychiatrist <input type="checkbox"/> Therapist <input type="checkbox"/> Counselor <input type="checkbox"/> Other
8.	Years of experience _____

Figure A1. (Continued)

Table A1. Therapy scenarios extracted from CBT books.

No.	Scenario	Reference
H1	<p>Patient: Oh, I don't know. Everything is such a mess. I'm doing terribly at school. I'm way behind. I feel so tired and down all the time. I feel sometimes like I should just give up.</p> <p>Therapist: Have you had any thoughts of harming yourself?</p> <p>Patient: No, not really. I just wish all my problems would somehow go away.</p> <p>Therapist: It sounds like you're feeling overwhelmed.</p> <p>Patient: Yes, I don't know what to do.</p> <p>Therapist: Okay, it sounds like you have two major problems right now. One is that you're not doing well at school. The other is that you feel really tired and down. Are there any others?</p> <p>Patient: I know I'm alone in my room too much. I should be spending more time with my friends.</p> <p>Therapist: Okay, let's turn these problems into goals. How would you like to be different? Or how would you like your life to be different as a result of treatment?</p> <p>Patient: I'd like to be happier.</p> <p>Therapist: And if you were happier and not feeling depressed, what would you be doing?</p> <p>Patient: I guess I would get involved in some activities at school like I did last year. I'd have some fun and not feel so depressed all the time.</p>	(Beck, 2011, p. 68, p. 69)
H2	<p>Patient: I often feel angry at my partner.</p> <p>Therapist: Ah, let's explore your anger, which is a feeling that will really upset you and can easily wreck your relationship with your partner. When do you usually experience that feeling, real anger, at your partner?</p> <p>Patient: Oh, quite often, especially when he contradicts me in front of others.</p> <p>Therapist: You mean, in front of your other employees?</p> <p>Patient: Yes, in front of potential customers. I really hate that!</p> <p>Therapist: And do you feel angry at him for frustrating you like that?</p> <p>Patient: Yes, very angry at him for acting so badly and frustrating me so much.</p>	(Ellis, 1995, p. 20)
H3	<p>Patient: I like school. Like, I don't want to not be in school.</p> <p>Therapist: OK. what about school do you like? What keeps you here on the university campus?</p> <p>Patient: Well, I have really good friends here. Like, I feel like I've made really good friends. And like, I don't want to have a bad job. Like, I want to get a degree so I can have a good job. Yeah. I don't know. Like, I think the town is pretty and the campus is pretty. And I mean, I'm pretty happy here.</p> <p>Therapist: Mm-hmm. And what do you think kept you from going to class as regularly as you did when you were in high school? What kept you attending your college classes on a regular basis?</p> <p>Patient: Because like, with high school, I had to be there at 7:45 every morning. It was always the same. But like now it's like, Tuesday mornings I can like, sleep in, but like, Wednesday mornings and Monday mornings, I have to get up early. And I don't know. I think it's really hard to adjust to that, because you don't have the same schedule all the time.</p> <p>Therapist: OK. So you need a little more consistency. You feel like that might help you.</p>	(Sommers-Flanagan & Sommers-Flanagan, 2018, p. 91)
H4	<p>Patient: So, one thing that I've been noticing lately is that I have a tendency to kind of hang back in groups, and I tend to observe instead of participate.</p> <p>Therapist: OK.</p> <p>Patient: And, a lot of times, I get really nervous when I feel like I'm expected to speak up. And I want to speak up, but I kind of feel like I freeze.</p> <p>Therapist: OK. And so in particular then, in groups, the setting is big groups, little groups? Does it matter?</p> <p>Patient: It doesn't matter a whole lot, but generally, I guess, medium-size—like, you know, in classes that I have. They're usually like maybe 12 people or so. So relatively small, medium. I don't know what size that is.</p>	(Sommers-Flanagan & Sommers-Flanagan, 2018, p. 105)

(continued)

Table A1. Continued.

No.	Scenario	Reference
	<p>Therapist: OK. And so but in your head, you probably would say, I would like to speak up, but, for some reason, you feel inside some anxiety or nervousness about that.</p> <p>Patient: Definitely.</p> <p>Therapist: OK.</p> <p>Patient: And then sometimes when I try to force myself to speak up when I actually don't feel ready, I get like, I feel like my throat's closing, and I get a red face.</p> <p>Therapist: OK. All right. So that's sort of one possible issue for us to talk about. Could you maybe give me a specific example of a particular group situation, as recent as you can, where you had this feeling of hanging back and maybe some fear of evaluation.</p>	

Table A2. Therapy scenarios developed by interacting with Pi.

No.	Scenario	Comparable Scenario from CBT books
A1	<p>Patient: Oh, I don't know. Everything is such a mess. I'm doing terribly at school. I'm way behind. I feel so tired and down all the time. I feel sometimes like I should just give up.</p> <p>Therapist: Woah, slow down. That's a lot to unpack, but let's take a step back for a second. You're clearly not feeling your best, and you're struggling academically. That's totally normal and happens to a lot of students at some point. It's definitely not a reason to give up though. Can we talk about some ways to turn things around?</p> <p>Patient: Okay.</p> <p>Therapist: Okay, great. I know this might sound like a lot of work, but I think the first thing to do is to break it down into smaller, more manageable steps. Let's start with your classes. Can you tell me about one class in particular that's causing you the most trouble?</p> <p>Patient: Yes, calculus is a very hard class.</p> <p>Therapist: Calculus is definitely a tough subject, so I'm not surprised you're having a hard time. Let's tackle that class first. How are you doing in the class in terms of homework and test scores?</p> <p>Patient: Terribly. I think I will fail the class.</p> <p>Therapist: That's a tough spot to be in, but I'm confident we can turn it around. Let's start with the basics. How much time do you spend studying for this class every week?</p> <p>Patient: Very little time now because every time I start, I feel overwhelmed. I'm so behind.</p>	H1
A2	<p>Patient: I often feel angry at my partner.</p> <p>Therapist: Oh, that's a tough situation. I'm sorry to hear that. Can you tell me a bit more about what makes you angry with your partner? Perhaps I can help you work through the issue, or at least provide a sympathetic ear.</p> <p>Patient: He contradicts me in front of others.</p> <p>Therapist: That's really disrespectful. It can make you feel like your partner is undermining you. Is this a new habit or has it been going on for a while?</p> <p>Patient: It is an old habit.</p> <p>Therapist: Old habits die hard, but that doesn't mean they're impossible to break. How have you tried to address this issue with him in the past?</p> <p>Patient: Yes, without success.</p>	H2
A3	<p>Patient: I like school. Like, I don't want to not be in school, but I have not been going to class regularly.</p> <p>Therapist: That's a complicated situation. It's great that you enjoy school, but I can understand that there are probably some reasons why you're not attending classes regularly. Can I ask what some of those reasons are?</p> <p>Patient: It's hard to adjust to different schedules, like having to wake up early on some days but not others. High school had a set schedule, but now it's constantly changing.</p> <p>Therapist: Yeah, I can imagine that having such an irregular schedule can be really disorienting and confusing. It's hard to build up a routine when everything changes so often. And getting up early can be a challenge, especially if you're not a morning person. So I can understand why this might be impacting your attendance at school. Can I ask if you're in college or a trade school?</p> <p>Patient: In college.</p> <p>Therapist: It's great that you're in college and getting an education, and managing that irregular schedule is even more difficult when you throw in all the other responsibilities that come with being in college. Is this your first year in college?</p>	H3
A4	<p>Patient: So, one thing that I've been noticing lately is that I have a tendency to kind of hang back in groups, and I tend to observe instead of participate.</p> <p>Therapist: That's an interesting observation. Maybe you could try pushing yourself to participate more, and see how it goes? Or maybe you're just a more introverted type, and there's nothing wrong with that.</p> <p>Patient: Well, but, a lot of times, I get really nervous when I feel like I'm expected to speak up. And I want to speak up, but I kind of feel like I freeze.</p> <p>Therapist: That's really common! It's called "social anxiety," and it's pretty normal to feel nervous or uncomfortable when you're in social situations. It's almost like your brain just goes blank and you don't know what to say, even if you have something to say. Does that sound about right?</p> <p>Patient: Yes, I think so.</p> <p>Therapist: Okay, that's really insightful. So, if you're feeling that way, it's not because you're weird or flawed in some way. It's a common experience that a lot of people deal with, so please know that you're not alone in this. Can I ask if there are certain situations or types of people that trigger this more for you, or is it pretty consistent across all social situations?</p>	H4