ISSN: 2584-0142
PSYCHOPEDIA JOURNALS
International Journal of Interdisciplinary Approaches in Psychology (IJIAP)

# Exploring Self-Disclosure in Conversations with AI Chatbots Vs. Human Mental Health Support among Young Adults

[1]**Tanvir Hoque**, *Kristu Jayanti Deemed to be University, Bangalore*
tanvirhq333@gmail.com
[2]**Ann Mary Jacob**, *Kristu Jayanti Deemed to be University, Bangalore*
ann.mj@kristujayanti.com

***Abstract:*** The increasing processes of artificial intelligence integration into mental health services result in the paradigm shift in the manner in which individuals seek emotional support and share their personal thoughts. Mental health chatbots powered by artificial intelligence provide anonymity, availability, and immediate conversation, either as an adjunct to human therapists or as an alternative to human therapists. However, there are no empirical studies on the process of self-disclosure in AI-mediated communication and traditional human therapy. In this qualitative study, the researchers examine the self-disclosure experiences of young adults when they interact with AI-powered mental health chatbots compared to human mental health professionals. It used a phenomenological research approach, and semi-structured interviews were conducted with the participants who had experience with both. To identify the patterns based on emotional experiences, depth of disclosure, and factors influencing the willingness of people to disclose, thematic analysis proposed by Braun and Clarke (2006) was employed for analyzing the data. The findings show a very clear set of differences between the two forms of relationship. The idea of AI chatbots was perceived as being very unpersonal and unemotional, a facilitator of superficial self-disclosure with very little emotional spilling. The main aim of the AI chatbots for the participants was to cope with stress, gain information, or demonstrate curiosity rather than emotional support. On the other hand, the relationships with human mental health professionals were characterized by warmth, empathy, validation, and psychological safety, which facilitated the disclosure of more and more authentic information about self, family, and emotional topics. The obstacles to self-

disclosure were stigma and privacy concerns, and empathy and non-judgmental response were the elements that helped to overcome these obstacles. Overall, the study illustrates the complementary role of AI chatbots in mental health and the inability of the latter to substitute human interaction as it contributes to meaningful self-disclosure.

**Keywords:** *Self-Disclosure, Artificial Intelligence Chatbots, Mental Health Practitioners, Empathy, Trust, Qualitative Study, Thematic Analysis, Psychological Support.*

## 1. Introduction

### *Background of the Study*

Self-disclosure is the process of revealing one's personal feelings, emotions, and thoughts to others. It is the basis of trust, intimacy, and effective relationships. One of the most important theories that explain this phenomenon is the Social Penetration Theory (Taylor & Altman, 1975, 1987). This theory explains relationship-building as a process of gradually moving from the outer, superficial layers to the inner, more intimate layers through mutual and rewarding self-disclosure. The "onion" model describes this process well: "As time passes, people reveal more and more about themselves, moving from public beliefs and feelings to increasingly private, inner thoughts."

Traditionally, self-disclosure occurred face-to-face. However, with the current rapid progress in digital technology, self-disclosure has overflowed into cyberspace. Online platforms provide anonymity, ease of access, and reduced fear of judgment, making people more likely to open up about their personal concerns—particularly those that carry a stigma or consequences in social relationships. In the past decade, AI-powered mental health chatbots have gained popularity as an emotional support system, providing instant comfort at any time and from anywhere. This fits well with the Online Disinhibition Effect, where online spaces decrease the inhibition to share personal information since the fear of being judged is reduced. Chatbots are perceived as non-judgmental and emotionally neutral, which may facilitate early self-disclosure (Widener & Lim, 2020; Lee & Lee, 2023).

However, despite these benefits, a lack of genuine empathy and emotional connection, which are crucial elements of therapeutic relationships and self-disclosure, exists in AI systems. Cultural factors also play a role. In Indian culture, emotional display is often inhibited because of mental health stigma, collectivist values, and social norms. However, increased acceptance of digital mental health resources among Indian youth provides new opportunities for self-disclosure. Young adulthood is a critical life stage for identity development and intimacy needs, making self-disclosure particularly important during this phase of life. Although AI chatbots

provide ease and anonymity, they cannot serve as a complete substitute for human mental health professionals, who provide critical ethical accountability, emotional nuance, and psychological safety.

Thus, the objective of this research is to investigate and compare self-disclosure experiences between AI chatbots and human mental health professionals among young adults in the Indian setting.

## 2. Theoretical Background

Social Penetration Theory highlights the importance of reciprocity, closeness, and vulnerability as we open up to others. However, when the conversation is handled by AI, true reciprocity is not actually taking place—it is only a simulation. This goes against traditional notions of intimacy, and one wonders whether disclosure is actually relational and therapeutic or merely instrumental when using AI. Variables such as empathy, trust, felt understanding, and individual differences in personality and levels of anxiety influence people's decisions about disclosure. In India, where mental health professionals are in short supply and waiting lists are long, the use of AI-based tools becomes even more attractive. Chatbots can be used as a gateway to mental health services, but the psychological effects of these tools are not yet understood. The current study takes a complementary approach: perhaps AI chatbots and human mental health professionals have different roles to play. AI can be used to help people articulate and express their emotions, and human mental health professionals can help facilitate more relational healing. Comparing the disclosure experiences on both platforms can help to develop theory and facilitate the integration of mental health services.

### 2.1 *Operational Definitions*

Self-Disclosure: The process of sharing one's thoughts, feelings, and experiences with others. In this study, it refers to the extent to which the disclosures made by the participants are detailed, straightforward, and emotionally nuanced.

Artificial Intelligence Chatbot Mental Health: Computer-based systems that offer mental health support through chat interfaces (e.g., Wysa, Woebot).

Human Mental Health Support: Psychological support offered by licensed professionals such as therapists and counselors, who are characterized by empathy, emotional understanding, and a commitment to ethics.

### 2.2 *Conclusion*

Self-disclosure is at the very heart of emotional healing and relationship-building. While AI may offer new ways of expressing oneself, it cannot substitute for the depth and

empathy of human relationships. A balanced and ethical approach is required to develop mental health support systems that integrate technology without diminishing human interaction.

## 3. Methodology

The methodology of this research is presented in a clear manner: it describes the problem, lists the objectives, formulates research questions, and explains the design, setting, participants, sampling procedure, tools and instruments, procedures, ethics, and data analysis. The methodological framework revolves around investigating and comparing how young adults experience self-disclosure during conversations with AI chatbots and when seeking support from human mental health professionals.

### 3.1 *Statement of the Problem*

With the growing development of artificial intelligence, there is a rising trend of using AI chatbots for emotional support and mental health-related conversations. Young adults tend to use these services for their convenience, anonymity, and non-judgmental nature. Self-disclosure is an essential component of psychological support and therapy, but there is a lack of qualitative information on how people disclose themselves to AI chatbots compared to disclosures to mental health professionals. There is a lack of in-depth knowledge about emotional comfort, perceived understanding, trust, and depth of disclosure in both situations. Thus, this research aims to explore how young adults disclose personal issues when interacting with AI chatbots versus human mental health professionals.

### 3.2 *Objectives of the Study*

1. Investigate the subjective experience of disclosing personal issues to AI chatbots.
2. Investigate the disclosure of personal issues to human mental health professionals.
3. Compare emotional comfort, perceived understanding, and depth of self-disclosure between the two.

### 3.3 *Research Questions*

1. What is the experience of self-disclosure for individuals when they engage in conversations with AI chatbots?
2. What are the descriptions of self-disclosure experiences when individuals receive human mental health support?
3. In what ways do emotional tone, comfort, and trust vary between self-disclosure experiences with AI chatbots and human mental health professionals?
4. What are the influences on an individual's willingness to self-disclose in each context?

### 3.4 *Research Design*

The research will use a qualitative phenomenological design to explore the experiences of individuals who have engaged with AI chatbots and human mental health professionals. Phenomenology is an appropriate methodology to use when exploring the meanings, emotions, and perceptions of self-disclosure, enabling individuals to describe their experiences in detail without losing the essence of their personal experiences.

### 3.5 *Research Setting*

The research will be conducted both online and offline. Participants will be recruited from social media platforms and educational institutions. Interviews will be conducted face-to-face or online, depending on the convenience of the participants.

### 3.6 *Participants*

Fifteen young adults aged 22 to 25, who had experience with chatbot AI conversational interfaces and human mental health professionals, participated. Participants were selected for their capacity to reflect upon and articulate their experiences with self-disclosure in either context.

### 3.7 *Sampling Technique*

A non-probability, purposive sampling technique was employed, which is common in qualitative research. Participants were selected according to criteria associated with the research rather than a random sampling technique.

### 3.8 *Sample Size*

Fifteen participants are adequate for phenomenological analysis, allowing for an in-depth examination of experiences.

### 3.9 *Tools and Instruments*

Semi-Structured Interview Guide. The interview guide was designed with open-ended questions to guide data collection while allowing participants to share their experiences freely. This approach allowed for the examination of themes such as emotional comfort, perceived understanding, trust, and the level of self-disclosure that emerged.

### 3.10 *Procedure*

Participants were informed about the nature and purpose of the research. Written informed consent was obtained. Interviews were conducted either face-to-face or online, audio-recorded with participants' consent, and transcribed verbatim for analysis.

### 3.11 *Ethical Considerations*

The research followed APA guidelines for ethical research. Participation was voluntary, anonymity and confidentiality were maintained, and participants were free to withdraw from

the research at any time without consequence.

### 3.12 *Method of Data Analysis*

***Thematic Analysis***

The data analysis was conducted using Braun and Clarke's (2006) thematic analysis. The thematic analysis consisted of six phases: familiarization with the data, coding, developing potential themes, reviewing the themes, defining the final themes, and then producing the report.

## 5. Results

### 5.1 *Introduction*

The current study reports the cognitive thematic analysis of young adults' experiences of self-disclosure in interactions with AI-based mental health chatbots and human mental health professionals. Participant Demographics: The participants were young adult postgraduate students aged 22-25 years, pursuing psychology-related courses. All participants were well aware of AI-based mental health chatbots and traditional human mental health professionals.

***Thematic Analysis***

The data resulted in seven major themes

**Theme 1:** *Emotional Neutrality and Mechanical Nature of AI Interactions*

| Sub-Theme | Description | Supporting Dialogues from Respondents |
|---|---|---|
| Emotional Neutrality | Participants felt emotionally neutral after interacting with AI chatbots, with no emotional release. | "I just felt neutral because it didn't feel like actual sharing." |
| Mechanical Responses | AI responses were perceived as robotic and scripted. | "It felt more like a device than a listener." |
| Lack of Warmth | Absence of emotional warmth reduced engagement. | "The responses felt pre-programmed rather than soulful." |
| No Sense of Being Understood | Participants felt misunderstood by AI. | "I didn't feel like I was understood." |
| Absence of Emotional Relief | AI failed to provide emotional relief. | "My emotions weren't released when I used Wysa." |

**Theme 2:** *Nature and Extent of Self-Disclosure with AI*

| Sub-Theme | Description | Supporting Dialogues from Respondents |
|---|---|---|
| Surface-Level Disclosure | Disclosure was very general. | "I would ask only very general questions." |
| Avoidance of Personal Problems | There was no disclosure of personal problems. | "I'm not ready to discuss my background." |
| Use of AI as Informational Resource | AI was used as an informational resource, not for emotional support. | "I just typed a few sentences about stress." |
| No Emotional Vulnerability | There was no emotional vulnerability. | "Chatbots are on the surface." |
| Curiosity-Driven Use | Use was curiosity-driven, not need-driven. | "I installed it just to get to know the AI." |

**Theme 3:** *Emotional Experience with Mental Health Professionals*

| Sub-Theme | Description | Supporting Dialogues from Respondents |
|---|---|---|
| Emotional Warmth | Mental health professionals showed warmth. | "The tone of my counselor felt very caring." |
| Feeling Heard | Mental health professionals allowed the participants to feel heard. | "She listened and gave me perspective." |
| Validation | Feeling validated increased comfort. | "I felt validated." |
| Emotional Safety | Therapy provided a safe environment. | "The therapist created a safe space." |
| Emotional Connection | An emotional connection was made. | "I felt emotionally connected." |

**Theme 4:** *Depth and Authenticity of Disclosure with Human Professionals*

| Sub-Theme | Description | Supporting Dialogues from Respondents |
|---|---|---|
| Deep Emotional Sharing | Disclosure was emotionally deep. | "I shared my most personal experiences." |
| Family and Personal Issues | Sensitive topics were discussed. | "I wouldn't have tried telling those to a chatbot." |
| Emotional Vulnerability | Participants allowed vulnerability. | "Therapy was about deep emotional confessions." |
| Greater Depth than AI | Disclosure was deeper than with AI. | Disclosure was deeper than with AI. |

| | | |
|---|---|---|
| Authentic Expression | Feelings were expressed honestly. | "The therapist really cared about me." |

**Theme 5:** *Trust and Psychological Safety*

| Sub-Theme | Description | Supporting Dialogues from Respondents |
|---|---|---|
| Trust in Human Professionals | The participants trusted human therapists. | "I trusted my therapist completely." |
| Lack of Trust in AI | The participants did not trust AI emotionally. | "I have no trust in the chatbot." |
| Non-Judgmental Attitude | Lack of judgment promoted trust. | "She never judged me." |
| Trust Developed Over Time | Trust developed over time. | "Once trust was established, it became easier." |
| Psychological Safety | Feeling safe promoted disclosure. | "I felt secure emotionally." |

**Theme 6:** *Barriers to Self-Disclosure*

| Sub-Theme | Description | Supporting Dialogues from Respondents |
|---|---|---|
| Fear of Judgment | Fear limited disclosure initially. | "Fear of judgment at first." |
| Cultural Stigma | Therapy-related stigma existed. | "Emotional sharing is taboo in our culture." |
| Privacy Concerns with AI | Data privacy reduced trust in AI. | "I didn't feel personal with AI." |
| Initial Hesitation | Opening up was difficult initially. | "It was really hard to open up." |
| Overcoming Barriers Through Empathy | Empathy reduced barriers. | "Empathy helped me overcome that." |

**Theme 7:** *AI as a Complementary Support System*

| Sub-Theme | Description | Supporting Dialogues from Respondents |
|---|---|---|
| Anonymity | AI facilitates anonymity. | "It is easy to converse with an AI because I can remain anonymous." |
| Accessibility | AI is easily accessible. | "AI can serve as an initial step." |
| Use in Low-Intensity Stress | AI is useful in low-intensity stress. | "I am able to discuss my stress." |

| | | |
|---|---|---|
| Need for Emotional Intelligence | AI lacks emotional intelligence. | "AI requires greater emotional intelligence." |
| AI Cannot Replace Humans | Human empathy is irreplaceable. | "Nothing can substitute for human empathy." |

## 6. Summary and Conclusion

### 6.1 *Findings*

The findings showed a marked difference in the nature of self-disclosure in AI-mediated communication and human mental health support. While AI chatbots were found to be emotionally neutral and information-oriented, human therapy was emotionally engaging and open. Although AI communication is anonymous and available, it is emotionally shallow and lacks trust. In contrast, human mental health practitioners enable progressive self-disclosure through genuine empathy, trust, and a feeling of psychological safety.

### 6.2 *Conclusion*

While AI chatbots are useful supplementary tools for mental health, they cannot substitute human mental health practitioners and their emotional integrity. Human empathy is still necessary for self-disclosure and psychological healing.

### 6.3 *Limitations*

The small sample size and qualitative research design limit generalizability. Psychology professionals in the sample may have biased their perceptions and responses.

### 6.4 *Future Research Suggestions*

Future research studies may use a mixed-methods design, larger and more representative samples, and cross-national comparisons. Hybrid AI-assisted therapy approaches may also be investigated.

## References

Beenish Moalla Chaudhry, & Happy Rani Debi. (2023). User perceptions and experiences of an AI-driven conversational agent for mental health support. *MHealth*, *0*. https://doi.org/10.21037/mhealth-23-55

Fitzpatrick, K. K., Darcy, A., & Vierhile, M. (2017a). 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

Ho, A., Hancock, J., & Miner, A. S. (2018). Psychological, Relational, and Emotional Effects of Self-Disclosure After Conversations With a Chatbot. *Journal of Communication*, *68*(4), 712–733. https://doi.org/10.1093/joc/jqy026

Inkster, B., Sarda, S., & Subramanian, V. (2018). An Empathy-Driven, Conversational Artificial Intelligence Agent (Wysa) for Digital Mental Well-Being: Real-World Data Evaluation Mixed-Methods Study. *JMIR MHealth and UHealth*, *6*(11), e12106. https://doi.org/10.2196/12106

Lee, J., & Lee, D. (2023). User perception and self-disclosure towards an AI psychotherapy chatbot according to the anthropomorphism of its profile picture. *Telematics and Informatics*, *85*, 102052–102052. https://doi.org/10.1016/j.tele.2023.102052

Warren-Smith, G., Laban, G., Pacheco, E.-M., & Cross, E. S. (2025). Knowledge cues to human origins facilitate self-disclosure during interactions with chatbots. *Computers in Human Behavior: Artificial Humans*, 100174. https://doi.org/10.1016/j.chbah.2025.100174

Wu, S. (n.d.). The effect of chatbots on self-disclosure among users. *ResearchGate* . https://doi.org/10.1201/9781032676043-91