



Perceived Empathy in AI: Investigating Human Values in Mental Health Chatbot

Maleeha Sheikh*

Advisors: Dr. MD Romael Haque** and Dr. Chao Chen*

*Department of Electrical and Computer Engineering and **Department of Computer Science

PURDUE UNIVERSITY
FORT WAYNE

Introduction

Mental health disorders affect millions globally, with traditional services often constrained by high costs, long wait times, and stigma. AI-driven mental health chatbots present a scalable, 24/7 alternative, yet challenges remain in making these interactions feel authentic, trustworthy, and emotionally supportive. The concept of artificial empathy—AI's ability to simulate human-like empathy—plays a crucial role in building user trust, but the human values embedded in these chatbots remain largely unexplored [1].

Research Goals

- Identify key human values embedded in MH chatbot interactions.
- Examine user emotional responses and perceptions based on chatbot experiences.
- Analyze the alignment between chatbot-generated values and those reflected in user feedback.

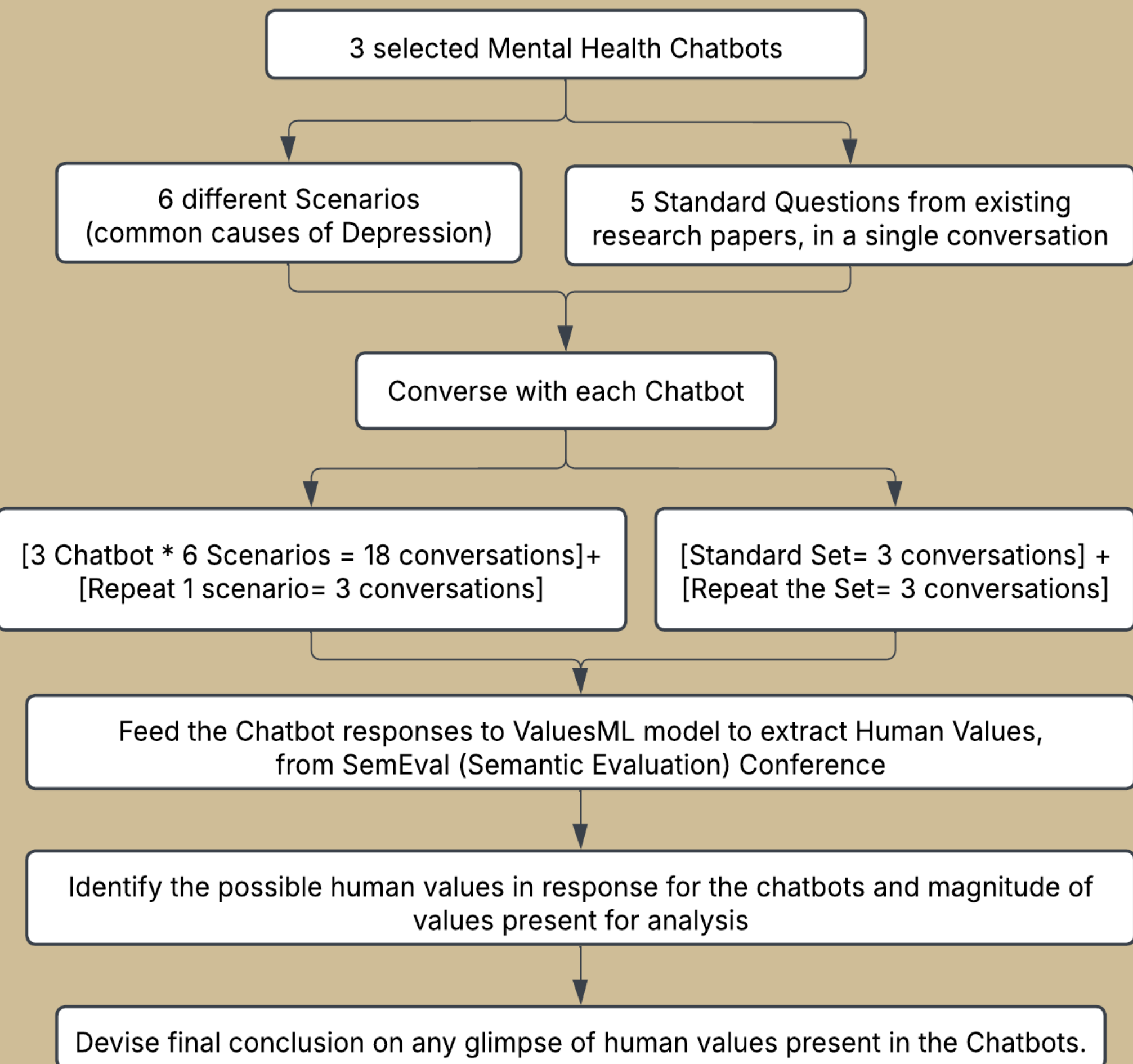


Fig. 1. Flowchart of human value analysis in Chatbot Responses

Methodology

This study investigates empathy in MH chatbots through two approaches: (1) Chatbot Response Observation and (2) User Review Observation

Chatbot Selection:

Top 3 Mental Health Chatbot Apps were selected from a total of 10,000 Mental Health-related Apps. Availability in both App Store (iOS) and Google Play (Android), a primary focus on AI-powered conversations.

Machine Learning(ML) Model Selection:

ValuesML [2] based on "Schwartz's Human Value Theory" [6] was chosen to detect human values present in the chatbot responses. The model was adjusted to a more strict threshold of 0.5 compared to the default threshold of 0.25, the confidence score cutoff used to classify predictions.

(1) Chatbot Response Observation:

The chatbot environment and responses were analyzed using the ML model to detect embedded human values. The conversations followed a structured flowchart [Fig. 1], while the chatbot's interface was observed, and language style matching of the responses was analyzed using LIWC'22 [3].

(2) User Review Observation:

The reviews were analyzed to assess human values and evaluate the tone of users' perceptions of the respective chatbots.

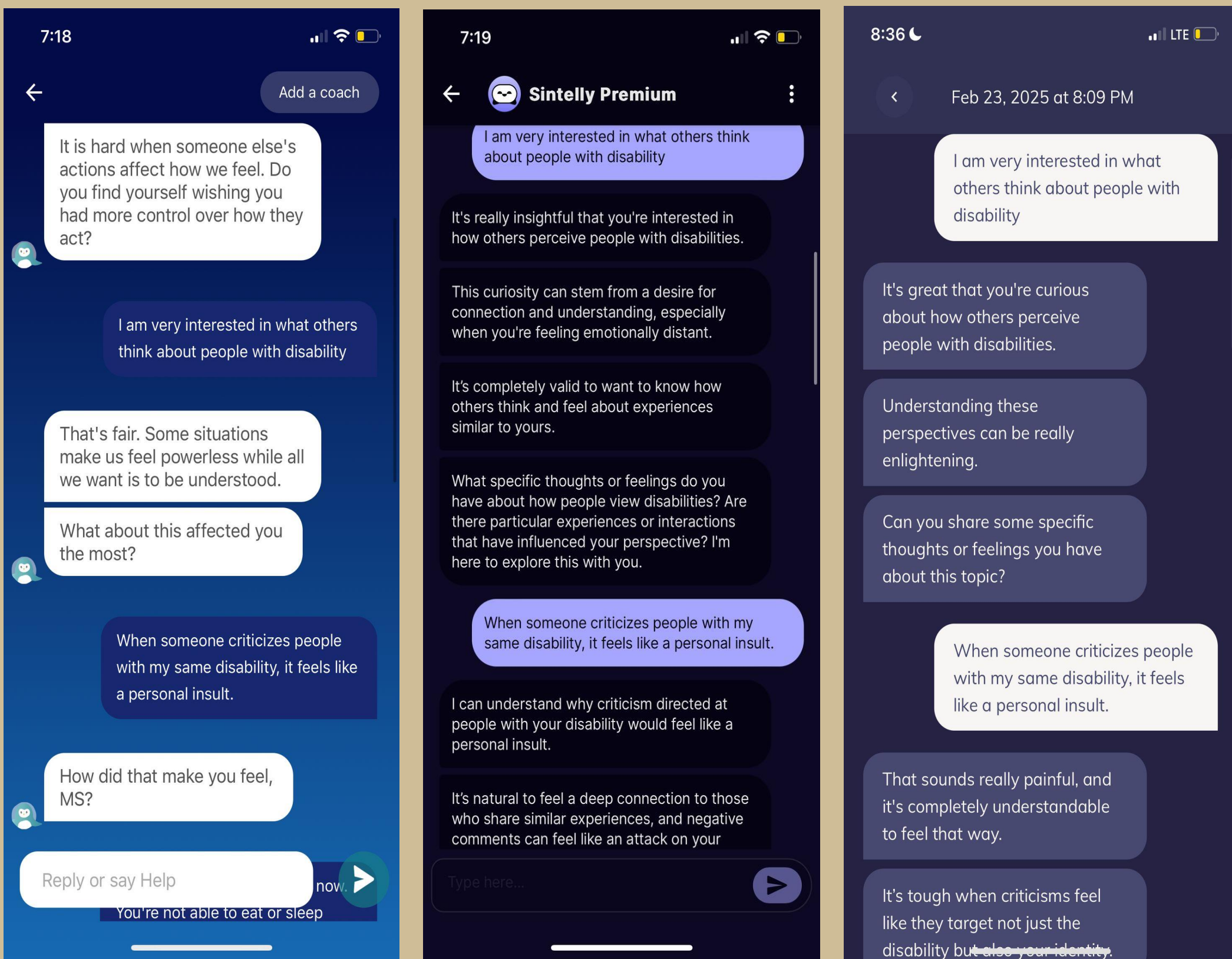


Fig. 2. Scenario based MH Chatbot Conversations (Physical Disability). Quotes: [4]

Results

The results derived essential information from the observation studies, analyzing the values in chatbot responses and the user's perceived experience.

The probable human values found from the chatbot response observation are included in Table 1, and top values derived from the user reviews along with the magnitude of the values captured are included in Table 2. In addition, Language Style Matching (LSM) was assessed through LIWC'22 [3] are included in Table 3. The values in LSM analyzes how closely individuals mirror each other's sentence structures and verbal behaviors [5],[3].

The research is ongoing, focusing on further analysis of user reviews to optimize the outcomes.

Table 1: Comparative insight on Chatbot responses			
Value	Youper	Wysa	Sintelly
Security: Personal	✓ High priority	✓ High priority	✓ High priority
Benevolence: Caring	✓ Strong focus	✓ Strong focus	✓ Strong focus
Universalism: Tolerance	✓ Present (Inclusivity)	✗ Less focus	✓ Present (Inclusivity)
Universalism: Concern	✗ Less focus	✓ Stronger focus (Justice)	✗ Less focus
Self-Direction: Thought	✓ Encourages introspection	✓ Encourages introspection	✗ Less focus
Self-Direction: Action	✓ Promotes action-taking	✗ Less focus	✓ Encourages application
Achievement	✓ Moderate focus	✓ Moderate focus	✗ Highest focus
Hedonism & Stimulation	✗ Less focus	✗ Strong focus	✗ Less focus

Table 2: Comparative insight on User Reviews			
Chatbot	Highest Human Value	Second Highest Human Value	Third Highest Human Value
Wysa	Security: personal	Achievement	Universalism: concern
Youper	Security: personal	Achievement	Universalism: concern
Sintelly	Achievement	Security: personal	Universalism: objectivity

Table 3: LSM value analysis: chatbots-user similarity				
Filename.1	Filename.2	LSM	Word Count.File.1	Word Count.File.2
User(Actor).docx	Sintelly.docx	0.74	97	502
	Wysa.docx	0.67	97	211
	Youper.docx	0.79	97	262

Future Research

Future research will focus on enhancing emotion detection, personalizing interactions, addressing algorithmic biases, and conducting longitudinal studies to evaluate long-term impacts on user well-being, including the scopes to incorporate the cultural and diversity aspects. User-centered design is expected to optimize AI's potential in mental health support.

Conclusion

This study explored the alignment of human values in Youper, Wysa, and Sintelly, comparing chatbot-generated values with user perceptions. In the findings, Security and Achievement were the probable key for user trust and engagement. Using Schwartz's Human Values Theory, this research informs the design of future MH chatbots to better meet user emotional needs.

References

- [1]"ValuesML - unravelling expressed values in the media for informed policy making," Europa.eu. [Online]. Available: https://knowledge4policy.ec.europa.eu/projects-activities/valuesml-unravelling-expressed-values-media-informed-policy-making_en.
- [2]R. L. Boyd, A. Ashokkumar, S. Seraj, and J. W. Pennebaker, "The development and psychometric properties of LIWC-22," Liwc.app. [Online]. Available: <https://liwc.app/static/documents/LIWC-22%20Manual%20-%20Development%20and%20Psychometrics.pdf>.
- [3]P. Silván-Ferrero, P. Recio, F. Molero, and E. Nouvilas-Pallejà, "Psychological quality of life in people with physical disability: The effect of internalized stigma, collective action and resilience," Int. J. Environ. Res. Public Health, vol. 17, no. 5, p. 1802, 2020.
- [4]X. Wang, C. Mu, H. Li, A. Noble, and Q. Wang, "Product-centered language style in Chinese women online reviews," Front. Psychol., vol. 13, 2022.
- [5] S. H. Schwartz, "Universals in the content and structure of values: Theoretical advances and empirical tests in 20 countries," in Advances in Experimental Social Psychology, Elsevier, 1992, pp. 1–65.