

Conversational AI for Mental Health: The Moderating Role of Big Five Personality Traits

Xuyuan Chen

*Faculty of Arts, University of British Columbia, Vancouver, Canada
arielchen718@gmail.com*

Abstract. Artificial intelligence is one of the most rapidly emerging research areas today. Researchers have found that, dumping one's troubles with AI could gain emotional support and moderate negative affects. However, there is still lack of research and summary on how personality factors influence the mental health intervention. Therefore, this paper examines the relationship between personality and the effects of AI conversation in the area of mental health by literature review. The research found that: (1) The interference of personality on the mental health intervention outcome has existed already. Therefore, studying the impact of personality differences on AI mental health intervention can be challenging, considering its' difficult to isolate the effect of users' current mental health status. (2) People with distinct personality hold different attitude about talking to AI and AI technology in general. For example, Neurotic users, on the one hand, stay vigilant on AI interaction; on the other hand, they wish AI could fulfill their social need of being understood and accepted. Based on the research, this article proposed two related recommendation to AI mental health product designing as following: (1) Considering that Individuals with certain personalities exhibit a higher need for the product, the priority and the designing focus should be given to the user group with personality traits that are targeted by the design features. (2) Future research should focus on the personality influences on the motivation and attention focus of AI chatbot using to improve the participation rate and the effect of mental health support.

Keywords: AI conversation, Personality, Mental Health

1. Introduction

AI technology has been progressively integrated into diverse everyday context, including but not limited to study, work, and daily interaction. With the advancement of modern technology, AI exhibits strong humanization in features like expression tone, emotion detection, expression consistency and so on [1]. Such progress in AI Chatbot design provides an almost authentic sense of companion. Moreover, such non-judgemental companionship is one of the core essential factors of effective treatment, considering it provides a safe environment for clients, facilitating the self realizations and emotional relief [2].

Nowadays, there are also more and more discussion arise about whether AI can replace the human mental health service and whether talking to AI is capable of alleviating distress. Although AI cannot replace diagnoses made by the professional psychiatrists, its capacity for providing

temporary companionship can address many psychological concerns in a cost-effective manner. Consequently, investigating the effectiveness of AI in psychological interventions and identifying the influencing factors of personality can provide valuable insight into this emerging area of research.

This article conducted a literature review and aim to, first, summarize the existing research on the therapeutic efficacy of AI dialogue for mental health. Then, propose how personality plays a role in AI mental health intervention. At the end, some suggestions will be made on the future AI mental health product design.

2. Key concept

2.1. AI conversation

Currently, there are many conversational AI agent on the market. Those agents are also known as AI chat bots. They can be approached based on mobile app, computers and other smart devices. A few examples would include ChatGPT, Gemini, Claude, DeepSeek and so on. This type of AI applications are capable of simulating humanized conversation. Given that interpersonal dialogue with real people has been demonstrated to alleviate negative emotions such as depression and anxiety [3], some users utilize these AI systems as "emotional buffers", and wish to mitigate short-term psychological distress. Such users may confide their emotional troubles, vent anger to the smart devices, or even just simply satisfy their need for self-expression. It is this specific form of interaction that is defined as "AI conversation" throughout this article.

2.2. Personality

This article will focus on the Big Five personality. The Big Five personality describes human personality through five core factors: openness, conscientiousness, extroversion, agreeableness, neuroticism. Its clear-cut taxonomy helps to explain the individual differences in acceptance of AI interventions in clearer manner. Openness is characterized by a preference for variety. Conscientiousness reflects a tendency of organization and diligence. Extroversion describes a person's sociability, and tendency to seek stimulation from the outside world. Agreeableness indicates a compassionate, cooperative, and trusting nature. Neuroticism measures emotional instability, anxiety, and susceptibility to stress. Together, these five dimensions provide a comprehensive and high-validity personality portrait for individuals [4,5].

3. AI Conversation and the effectiveness of mental health support

Given the novelty of simulated dialogue technology, most studies published within recent years, with the overall trend supports that communicating with AI is an effective method of alleviating negative affects..

He and colleagues summarized a wide range of research about conversational AI agent, including AI chat bot, VR conversational agent, embodied conversational agent and avatar therapy [6]. Those agent are not pre-set encoding codes, but AI robot that generates new conversations based on users' distinct feedback. Research proposed that, conversational AI agent interaction produced a statistically significant improvement in various perspectives of participants' mental health including depressive symptoms, generalized anxiety symptoms, specific anxiety symptoms, quality of life or well-being, general distress, stress, mental disorder symptoms, psychosomatic disease symptoms, and negative affect. However, due to various restriction, the long-term effect still needs to be confirmed.

Furthermore, another research that conducted among a university student population, compared the AI conversation with the traditional emotional expression [7]. The research found that, comparing to writing emotional diary, talking to AI was associated with a greater improvement in depressive symptoms. The finding illustrates that, trained AI chat bots do not merely provide an outlet for self-healing, but also offers structured support that enhances their clinical potential.

Cross-culturally, the effect still exists in studies where participants came from non-western background. In a research facing Chinese university students finds that CBT-based AI chatbot can significantly reduce depression and loneliness in Chinese university students. At the same time, this research also proposed that the effect is stronger effects among those experiencing high financial stress comparing to control group [8]. Moreover, financial stress is one of the most influential factors in mental health status, considering it is associated with persistent feeling of concern and helplessness. In addition to that, professional human psychology therapies are relatively expensive in general, which may impose additional strain on their already unstable living conditions. Consequently, financial stress further undermines their ability to seek and accept help. From this perspective, AI conversation is significantly more affordable than trained professionals in terms of costs. This method can act as a temporary buffer and assistant for people who are struggled in stress and financial difficulties.

Another research had found that, the extent to which communication with AI reduces negative affects is determined by the frequency of the interaction [9]. The effect remains statistically significant regardless of participants' baseline levels of their original depression level. Although improvements in the depression and anxiety are observed in participants with both high and low baseline levels of psychological well-being, the individuals with more severe symptoms show more progress. Additionally, the interaction frequency emerges as a stronger predictor than the symptoms severity when examining the effectiveness of AI conversation on mental health improvement: the more frequent the interactions participants had, the more depression and anxiety reduction will be made. Frequent interaction facilitates the establishment of a stable emotional regulation rhythm.

Research proposed that, comparing to the traditional process of self-enhancement via knowledge absorption, digital interactive treatment method possess several benefits [10]: (1) Accessibility: AI conversations are more available cross time and geographic constrains. Typing on a smart device does not require obtaining materials or making appointment with other parties in advance. (2) Empathy and individuality: AI conversation is capable of generating individualized responses tailored to the client's presenting issue. It's easier to build connections, feel supported and accept the information than the abstract and complex theories. (3) Interactive and engagement features: generally, users perceive AI conversations as more interesting and engaging because they offer an engaging and hands-on interactive experience. In summary, AI conversation enhances users' interaction willingness of emotion expression because its convenient, personalized, and interactive features, thereby better serving the field of mental health support.

In the meantime, drawback of AI conversation also exists. Conversation with a non-human being has limitations in several interactive functions, such as unnatural dialogue tone, inability to understand complex answers, and unexpected responses. Problems like dehumanized responses, rigid in feedback patterns, content repetition, overly vague or simplistic answers will reduce the participation rate. As a result, the interaction limitations may decrease user engagement because of the specific interaction details.

4. Personality factors on the efficacy of AI mental health interactions

Personality itself plays a role in affecting the outcomes of mental health interventions. Firstly, different personalities hold different attitudes towards seeking counseling or other professional support. For example, people with high openness are more willing to explore distinct helping methods when facing troubles, as a result, they are more likely to seek professional counselling help when facing negative affects like anxiety or depression [11]. Secondly, because individuals with certain traits are more predisposed to psychological disorders [12], they would have a stronger need for mental health support. Based on the above, personality factors influence both the willingness of mental health support seeking and psychological well-being, which makes the personality influences on AI agent intervention effects harder to be evaluated.

Furthermore, different personalities hold different attitudes toward AI technology and AI conversation. A study conducted on middle school students found that neuroticism level is positively correlated with the antipathy toward AI tools (eg. feel irritable and frustrated during interaction), and fear and anxiety towards the AI technology (eg. concerned that AI will replace them in the workforce) [13]. As mentioned in the previous section, participation rate is one of the most important factors in the effect of AI conversation. Such negative attitudes toward AI might decrease their participation rate and lead to ineffective result.

Meanwhile, humans' attitudes are complex: As Neurotic users possess more AI anxiety, they also wish that conversational AI agents could fulfill their social needs. It means that, Those with high neuroticism appear to view AI as warm and believe that the use of AI would satisfy their need for relatedness, although they may not necessarily believe that AI is functionally useful [14]. Therefore, those complex attitude might lead to different subjects. For example, neurotic people might more resist to talk about the technical details in the workplace with AI because of their anxiety and fear, but when it comes to interpersonal problems, they would disclose more emotional experiences and expect supportive feedback. For conscious people, when they would conceptualize AI tools into their work tasks, their participation rate will increase [15]. Thus, different personalities can lead to different points of attention to AI tool using, as well as different motivations for using AI, thereby affecting the frequency and effectiveness of AI mental health interventions.

Therefore, when designing AI chatbot for mental health support, the designer can modify the interactive features depending on the different personality characteristics to increase their participation frequency and acceptance of the operational mode of the AI agents. For example, based on the research of the characteristics of Neuroticism users, appropriate measures should be taken, such as clear privacy protection prompts, to reduce their hostility and anxiety. It can be explicitly and specifically stated that, their data will not be used to train other big data or used in a way that is harmful to them before the interaction. As for people who are high in consciousness, AI conversation can be reframed as a daily working task to increase usage frequency.

5. Discussion and suggestion

The current body of research indicates that AI conversation improve negative emotional states. Therefore, developing an AI chatbot that specifically focusing on providing mental support and assisting stress releasing is feasible and research-worthy. Meanwhile, future research needs to look into the different user characteristics and its interaction features when it comes to AI conversation.

From the perspectives of the Big Five personality, there are many differences between different personalities when it comes to how AI communicates. Developers first need to consider the different proportions of users. Some personality groups have a high proportion of individuals with poor

psychological conditions, so it is necessary to conduct more research on them when developing software, as the number or frequency of people using these personalities may be larger. Secondly, future research can focus on the motivations and areas of interest of individuals with different personalities when using AI chatbots, in order to better package and re brand features, enhance user satisfaction, and improve psychological support effectiveness. For example, Neurotic people would use AI to satisfy their needs for relatedness, but due to a lack of trust in information processing and unknown technologies, they would prioritize communicating with AI on technical matters similar to work. To solve this problem, developers can think more about how to ensure information security and gain users' trust.

6. Conclusion

This study first summarizes the effectiveness of AI dialogue as a psychological intervention to alleviate negative emotions in humans, and specifically highlights that frequency of use is an important factor affecting intervention outcomes; Subsequently, this paper discussed how personality factors affect the effectiveness of AI interventions on mental health and made two findings: (1) Individuals possessing personality traits strongly associated with a low desire for mental health support and positive mental health status, compared to those with a high need for mental health support and poor psychological well-being, the likelihood of requiring AI psychological intervention is lower. (2) People with different personalities have complex attitudes towards AI; They may have an attitude towards AI technology itself, but when using AI, they have another need. This article suggests in the field of AI intervention, future research should place more focus on study personality traits that have a greater demand for AI mental health support. Additionally, future research could study the motivation and attention focus of people with different personalities when interacting with AI chatbot agents to better understanding the user logic.

References

- [1] Zhang, Y., Yang, X., Xu, X., Gao, Z., Huang, Y., Mu, S., ... & Yu, G. (2024). Affective computing in the era of large language models: A survey from the NLP perspective. arXiv preprint arXiv: 2408.04638.
- [2] Rogers, C. R. (1951). *Client-centered therapy: Its current practice, implications, and theory*. Houghton Mifflin.
- [3] Harandi, T. F., Taghinasab, M. M., & Nayeri, T. D. (2017). The correlation of social support with mental health: A meta-analysis. *Electronic Physician*, 9(9), 5212–5222. <https://doi.org/10.19082/5212>
- [4] Tupes, E. C., & Christal, R. E. (1992). Recurrent personality factors based on trait ratings. *Journal of Personality*, 60(2), 225–251.
- [5] Norman, W. T. (1963). Toward an adequate taxonomy of personality attributes: Replicated factor structure in peer nomination personality ratings. *The Journal of Abnormal and Social Psychology*, 66(6), 574–583.
- [6] He, Y., Yang, L., Qian, C., Li, T., Su, Z., Zhang, Q., & Hou, X. (2023). Conversational agent interventions for mental health problems: Systematic review and meta-analysis of randomized controlled trials. *Journal of Medical Internet Research*, 25, e43862.
- [7] Liu, H., Peng, H., Song, X., Xu, C., & Zhang, M. (2022). Using AI chatbots to provide self-help depression interventions for university students: A randomized trial of effectiveness. *Internet Interventions*, 27, 100495.
- [8] Wang, Y., Li, X., Zhang, Q., Yeung, D., & Wu, Y. (2025). Effect of a cognitive behavioral therapy–based AI chatbot on depression and loneliness in Chinese university students: Randomized controlled trial with financial stress moderation. *JMIR mHealth and uHealth*, 13, e63806.
- [9] Karkosz, S., Szymański, R., Sanna, K., & Michałowski, J. (2024). Effectiveness of a web-based and mobile therapy chatbot on anxiety and depressive symptoms in subclinical young adults: Randomized controlled trial. *JMIR Formative Research*, 8(1), e47960.
- [10] Gual-Montolio, P., Jaén, I., Martínez-Borba, V., Castilla, D., & Suso-Ribera, C. (2022). Using artificial intelligence to enhance ongoing psychological interventions for emotional problems in real- or close to real-time: A systematic review. *International Journal of Environmental Research and Public Health*, 19(13), 7737.

- [11] Çekici, F. (2019). The relationships between the Big Five personality traits and attitudes towards seeking professional psychological help in mental health counselor candidates: Mediating effect of cognitive flexibility. *Educational Research and Reviews*, 14(14), 501–511.
- [12] Kang, W., Steffens, F., Pineda, S., Widuch, K., & Malvaso, A. (2023). Personality traits and dimensions of mental health. *Scientific Reports*, 13(1), 7091. <https://doi.org/10.1038/s41598-023-33996-1>
- [13] Stănescu, D. F., & Romaşcanu, M. C. (2024). The influence of AI anxiety and neuroticism on attitudes toward artificial intelligence. *European Journal of Sustainable Development*, 13(4), 191–191.
- [14] Park, J., & Woo, S. E. (2022). Who likes artificial intelligence? Personality predictors of attitudes toward artificial intelligence. *The Journal of Psychology*, 156(1), 68–94.
- [15] Arpacı, I., Kuşci, I., & Gibreel, O. (2025). The role of personality traits in predicting educational use of generative AI in higher education. *Scientific Reports*, 15(1), 30440.