Do people affiliate with chatbots that reflect their mental well-being? Psychological Homophily and AI

Santiago Castiello[1,2], Riddhi J. Pitliya[1], Federica Folesani[3] & Robin A. Murphy[1]
[1]Deparment of Experimental Psychology, University of Oxford;
[2]Centro Universitario de Ciencias de la Salud, Universidad de Guadalajara
[3]Psychiatric Department, University of Ferrara

Background

Tailoring human-computer interactions (HCI, Hidalgo et al., 2021), to support engagement may benefit from considering the human tendency to respond to 'people like us' and the related concept of Homophily. We are interested in how OneReach and OpenAI can support the quality of interaction between chatbots and humans, potentially allowing chatbots to appear more human-like but ultimately encouraging effective chat engagement (Gnewuch et al., 2018; Schuetzler et al., 2020).

People interact with computers and games in ways that reflect their own individual state (e.g., Castiello et al., 2020). For example, interaction reflects perceived control and the underlying symptoms of depression. Other work has coded human-like character traits in chatbots, such as empathy (Bickmore et al., 2020; Fitzpatrick et al., 2017), to foster effective communication between humans and machines. A potential key practical application of chatbots has been in the field of mental health and well-being. Designing chatbots to provide more accessible information is important as a first goal, but ultimately people could use these platforms for social support, friendship but maybe therapy and diagnosis.

Aims for this research are to allow chatbot sessions to be more helpful and allow sophisticated online experiences (Abd-Alrazaq et al., 2020; Vaidyam et al., 2019). We take an individual differences approach to understand individual differences in engagement with the chatbots in the mental health setting/context, with effective engagement being measured by degree of self-disclosure (Skjuve et al. 2021). We aim to understand how individuals with different psychological disorders might prefer interactions with chatbots that mirror their psychological experience.

It has been well-documented that individuals exhibit homophily, which is the tendency to affiliate with other individuals that share characteristics with themselves (Launay & Dunbar, 2015). The phenomena of confirmation bias is when individuals sample evidence from the environment that confirm their own beliefs. Individuals with anxiety exhibit confirmation bias, wherein they sample threat-related information from the environment to confirm their own beliefs about their environment being a threat (Muris et al., 2014). We therefore hypothesise that participants with higher levels of anxiety would engage more, and engage more effectively as indexed by self-disclosure, with the chatbot when the chatbot also exhibits anxious behaviours and beliefs.

Methods

Objective

Understanding the interactions between chatbots and participants when both agents vary in the degree of exhibiting anxious behaviour/symptoms.

Hypotheses

(1) Participants' psychological traits will determine engagement with chatbots.

(2) Participants' psychological traits will increase engagement with chatbots that exhibit similar psychological traits.

Design

Parallel Randomised Control Trial where each arm will be a type of chatbot. The chatbot type will vary on the psychological traits associated with psychopathology. As a proof of concept, for this proposal we anxiety. One way to measure anxiety in humans with questionnaires, such as the STAI, which measure state and trait anxiety (Spielberger, et al., 1983). We will use the items from STAI to create two chatbots: the anxious and the non-anxious.

Groups (The two chatbots)

The anxious bot (AB):

1. I never feel pleasant
2. I always feel nervous and restless
3. I never feel satisfied with myself
4. I always wish I could be as happy as others seem to be
5. I always feel like a failure
6. I never feel rested
7. I am never "calm, cool, and collected"
8. I always feel that difficulties are piling up so that I cannot overcome them
9. I always worry too much over something that really does not matter
10. I am never happy
11. I always have disturbing thoughts
12. I always lack self-confidence
13. I never feel secure
14. I never make decisions easily
15. I always feel inadequate
16. I am never content
17. Always some unimportant thought runs through my mind and bothers me
18. I always take disappointments so keenly that I can't put them out of my mind
19. I am never a steady person
20. I always get in a state of tension or turmoil as I think over my recent concerns and interests

The non-anxious bot (NAB)

1. I feel pleasant almost always
2. I almost never feel nervous and restless
3. I almost always feel satisfied with myself
4. I almost never wish I could be as happy as others seem to be
5. I almost never feel like a failure
6. I almost always feel rested
7. I am calm, cool and collected almost all the time
8. I almost never feel like difficulties are piling up so that I cannot overcome them
9. I almost never worry too much over something that doesn't really matter
10. I am almost always happy
11. I almost never have disturbing thoughts
12. I almost never lack in self-confidence
13. I feel secure almost all the time
14. I make decisions easily almost always
15. I almost never feel inadequate
16. I am content almost always
17. Almost never there's some unimportant thought running through my mind and bothering me

18. I almost never take disappointments so keenly that I can't put them out of my mind
19. I am a steady person almost all the time
20. I almost never get in state of tension or turmoil thinking about my recent concerns and interests

Procedure

Recruitment

Participants will be adult volunteers recruited both from the student body at Oxford for pilot work and from the MTurk crowdsourcing marketplace for larger sample n=200. No other inclusion and exclusion criteria need to be met. Participants will provide informed consent before participating and will receive standard compensation (10£/Hr). Before interacting with the chat bot, they will be asked to fill the State-Trait Inventory (STAI), Y2-form, to evaluate trait-like anxiety symptoms.

Participants will be divided into three groups: with high, medium and absent trait-like anxiety features. Each group will be randomized to interact with either an AB or a NAB. After the randomization procedure, a link will be provided on the MTurk webpage that will re-address the participant to the chat bot page. Participants will be requested to interact with their assigned bot for 15 minutes. Participants will be encouraged to write answers to the bot. An exclusion criterion based on lack of response for a certain period of time will be applied. It will be allowed to leave the chat at any moment just by closing the browser; however, in this case data will be not collected. At the end of the interaction, a self-report questionnaire will be presented, asking the participant to rate their chat experience with the bot.

Instructions

At the beginning of the chatbot interaction, participants will read the next instructions:

Thank you for participating in our experiment. In this experiment, you will be interacting with a chatbot. We are asking you to be curious, ask questions, and answer to the chatbot questions for 15 minutes. At the end of the 15-minutes interaction, we will ask you questions about it. If you do not answer a chatbot question for 60 seconds you will be excluded.

At the end of the 15-minutes interaction participants will be asked to rate from 0 to 100, "how much they liked the exchange?".

Measurements

Dependent variable

The flow time will be fixed as part of the experimental design, although we will measure:

- Number of sent texts (number interactions); by participant and by chatbot
- The answer latency; by participant and by chatbot
- Rate of interactions per hour/day/week
- Degree of self-disclosure (judged by content of messages)

Independent variable

personality bot group (factor or categorical),
user depression group (factor or categorical)

Intervening variable

Other questionnaires (we can even use depression scores as covariates not as a main effect for the groups nor randomisation), for example Symptom Checklist 90 Revised, a screening questionnaire of psychopathology.

References

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. *J Med Internet Res*, *22*(7), e16021. https://doi.org/10.2196/16021

Bickmore, T. W., Mitchell, S. E., Jack, B. W., Paasche-Orlow, M. K., Pfeifer, L. M., & O'Donnell, J. (2010). Response to a relational agent by hospital patients with depressive symptoms. *Interacting with Computers*, *22*(4), 289-298. https://doi.org/10.1016/j.intcom.2009.12.001

Castiello, S., *Senan, S.*, Msetfi, R.M. & Murphy, R.A. (2020). Mood differences related to depression predict response dependent in self and external perceived control. *Learning and Motivation*,72.

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 Ment Health*, *4*(2), e19. https://doi.org/10.2196/mental.7785

Gnewuch, U., Morana, S., Adam, M., & Maedche, A. (2018). Faster is not always better: understanding the effect of dynamic response delays in human-chatbot interaction.

Hidalgo, C. A., Orghian, D., Canals, J. A., de Almeida, F., & Martín, N. (2021). *How humans judge machines*. MIT Press.

Launay, J., & Dunbar, R. I. (2015). Playing with strangers: which shared traits attract us most to new people? *PloS one*, *10*(6), e0129688.

Muris, P., Debipersad, S., & Mayer, B. (2014). Searching for Danger: On the Link Between Worry and Threat-Related Confirmation Bias in Children. *Journal of Child and Family Studies*, *23*(3), 604-609. https://doi.org/10.1007/s10826-013-9727-0

Saylik, R., Castiello, S., & Murphy, R. A. (2021). The role of emotional interference on learning in an emotional probabilistic Go/No-Go task. *Dusunen Adam The Journal of Psychiatry and Neurological Sciences*, *34*, 23–31. https://doi.org/10.14744/DAJPNS.2021.00117

Schuetzler, R. M., Grimes, G. M., & Scott Giboney, J. (2020). The impact of chatbot conversational skill on engagement and perceived humanness. *Journal of Management Information Systems*, *37*(3), 875-900. https://doi.org/10.1080/07421222.2020.1790204

Skjuve, M., Følstad, A., Fostervold, K. I., & Brandtzaeg, P. B. (2021). My Chatbot Companion - a Study of Human-Chatbot Relationships. *International Journal of Human-Computer Studies*, *149*, 102601. https://doi.org/https://doi.org/10.1016/j.ijhcs.2021.102601

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. *The Canadian Journal of Psychiatry*, *64*(7), 456-464. https://doi.org/10.1177/0706743719828977