

# **Beyond Gender Stereotypes: Evaluating Trust and Value in Well-Being Chatbot Interactions. Extended Abstract**

Agnès HELME-GUIZON, Full Professor, Univ. Grenoble Alpes, Grenoble INP, CERAG, France  
[agnes.helme-guizon@univ-grenoble-alpes.fr](mailto:agnes.helme-guizon@univ-grenoble-alpes.fr)

Jade BROYER, Senior Market Researcher, Enov, France  
[jade.broyer@gmail.com](mailto:jade.broyer@gmail.com)

Soffien BATAOUI, Full Professor, Université Lyon 3, Magellan, France  
[soffien.bataoui@univ-lyon3.fr](mailto:soffien.bataoui@univ-lyon3.fr)

Mohammed Oualid HAKIMI, University of Prince Mugrin, Arabie Saoudite  
[m.hakimi@upm.edu.sa](mailto:m.hakimi@upm.edu.sa)

Keywords: chatbot; well-being; warmth; competence; gender stereotypes

## **Introduction**

Healthcare chatbots offer several advantages, such as 24/7 accessibility (Inkster et al., 2018), reduced stigma, lower cost (Lin et al., 2023), and effectiveness (Park et al., 2022). However, there is resistance due to a perceived lack of humanity. Humanisation through gender, names, and avatars can help address this issue. Research has shown that gender influences the intention to use chatbots (Borau et al., 2021). According to the CASA paradigm and Stereotype Content Model (Fiske et al., 2007), male chatbots are seen as assertive leaders, while female chatbots are perceived as more human, attentive (Borau et al., 2021), cooperative, and collaborative (Hentschel et al., 2019). Perceived warmth and competence enhance interaction value and encourage continued use (Belanche et al., 2021). Hall and Crisp (2005) suggest that traditional views can be reversed, and further validation is needed to understand counter-stereotype effects in healthcare contexts (Schillaci et al., 2024). Gaps remain in our understanding of how gender perceptions affect warmth, competence, trust, and value in well-being applications. Our study explores the impact of chatbot gender on users' perceptions of warmth, competence, trust, perceived value, and intention to use in two experimental settings: scenario-based and real-time interactions with a chatbot specifically designed by the researchers.

### **The effects of a chatbot's gender: Literature review and hypothesis**

Given the importance of human-like qualities such as warmth and competence for chatbot acceptance in mental health applications (Fitzpatrick et al., 2017), and that both warmth and competence are needed, we propose that female- and male-gendered chatbots will be perceived as equally warm and competent (Schillaci et al., 2024). Furthermore, as trust plays a critical role in human-chatbot relationships and is influenced by factors including chatbot competence, accuracy in understanding user requests, ability to provide useful responses, and politeness (Yen and Chiang, 2021; Følstad et al., 2018), we propose that higher perceptions of warmth and competence will positively affect trust in chatbots. This increased trust is expected to lead to favorable behavioral intentions such as positive attitudes and purchase intentions (Yen and Chiang, 2021; Wang et al., 2007). Drawing from research on robots that distinguishes functional, emotional, and social values (Belanche et al., 2021), we further hypothesize that competence will more strongly influence utilitarian values, while warmth will have a greater effect on emotional value, with both dimensions jointly impacting users' attitudes toward and intention to use chatbots.

### **Study 1 - Experimental protocol, data collection and analysis**

*Objective.* Study 1 aimed to investigate how the gender of a chatbot influences users' perception of warmth, competence, trust, and intention to use in a scenario-based experimental setting

*Procedure.* A total of 297 prolific participants were randomly assigned to male- or female-gendered chatbot conditions. Gender was manipulated using names from Borau et al.(2021) and gender-normative avatars (see Appendix 1). Pre-testing confirmed the successful gender assignment. After viewing a chatbot screenshot with a welcome message, participants reported their trust, attitude, usage intention, and evaluated the chatbot's warmth and competence.

*Measures.* Warmth and competence were measured using Eyssel and Hegel's (2012) adapted scale. Trust was measured using Pitardi and Mariott (2021), attitude following Borau et al. (2021), and behavioral intentions using Liu and Tao's (2022) adapted scale. A gender manipulation check verified the correct gender assignment. All items were measured using a 5-point Likert scale.

*Data Analysis.* Factor scores were calculated for each scale. Anovas and regression analyses validated our assumptions

## **Study 2 - Experimental protocol, data collection and analysis**

*Objective.* This study replicated Study 1 using real-time interactions with a well-being chatbot.

*Procedure.* A website with a researcher-designed Tidio chatbot was created to avoid risky response patterns (De Freitas et al., 2023). Participants were randomly assigned to male or female-gendered chatbots and interacted freely, receiving exercise recommendations. All other chatbot features were identical. A total of 474 participants completed the survey after providing consent. Following the interaction, they reported perceived warmth, competence, trust, value, attitude, and usage intention.

*Measures.* Similar to Study 1, value was assessed using Belanche et al. (2021). A manipulation check confirmed accurate chatbot gender identification. All the scales used a 5-point Likert format.

*Data Analysis.* Factor scores were calculated for each scale. Anova and regression analyses validated our assumptions

## **Discussion, limitations and research avenues**

Contrary to research suggesting that female-gendered AI agents evoke higher perceptions of warmth, the findings, consistent with Schillaci et al. (2024), revealed no significant differences between male- and female-gendered well-being chatbots in terms of user attitudes, intentions, or perceptions of warmth and competence. This challenges the notion that gender stereotypes influence chatbot interactions in the context of well-being. Warmth and competence positively affected trust, which emerged as a crucial mediator in shaping user attitudes and intentions to use the chatbot. Both warmth and competence positively influenced functional and emotional values. Emotional value had a stronger impact on user attitudes and perceived well-being than functional value, echoing the findings of healthcare chatbot studies. This underscores the importance of designing AI systems that prioritise emotional resonance in mental health. While the gender of the chatbot may not directly influence perceptions, the design should focus on building trust through warmth and competence, with emotional value fostering positive user experiences.

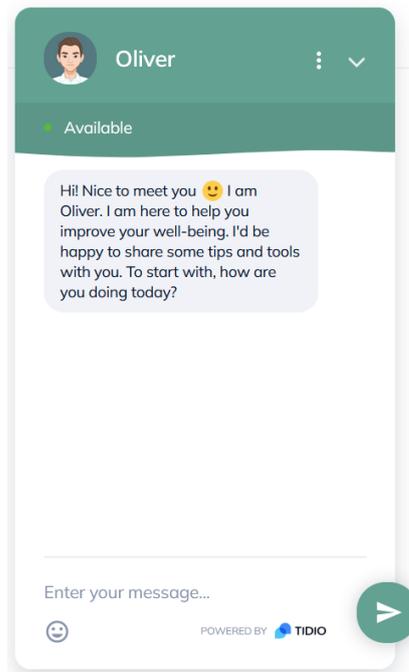
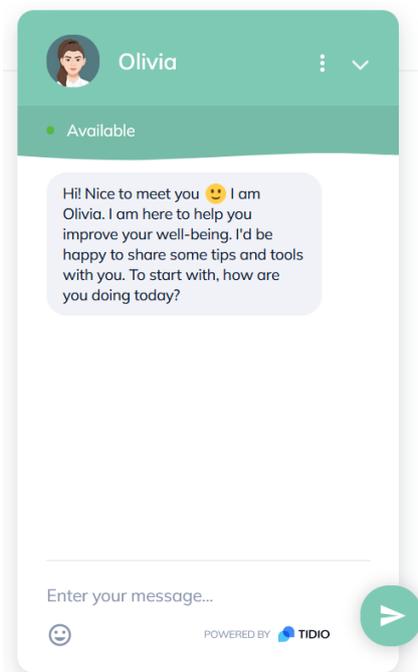
*Limitations and future research directions* This study manipulated the gender of chatbots through names and avatars. Future studies could strengthen this through voice and language styles to enhance gender stereotypes. Findings on counterstereotypes regarding warmth and competence are more reliable (Schillaci et al., 2024). Gender congruence between chatbots and users could be investigated (Pitardi et al., 2023; Zogaj et al., 2023).

This study advances theoretical understanding and practical applications by examining how gender influences perceptions of warmth and competence in chatbot interactions, extending the research beyond commercial contexts. This study provides recommendations for implementing chatbot gender and encourages a critical analysis of gender in technology, particularly regarding societal stereotypes.

## References

- Belanche, D., Casaló, L.V., Schepers, J., & Flavián, C. (2021). Examining the effects of robots' physical appearance, warmth, and competence in frontline services: The Humanness-Value-Loyalty model. *Psychology & Marketing*, 38(12), 2357–2376.
- Borau, S., Otterbring, T., Laporte, S., & Fosso Wamba, S. (2021). The most human bot: Female gendering increases humanness perceptions of bots and acceptance of AI. *Psychology & Marketing*, 38(7), 1052-1068.
- De Freitas, J., Uğuralp, A.K., Oğuz-Uğuralp, Z., & Puntoni, S. (2024). Chatbots and mental health: Insights into the safety of generative AI. *Journal of Consumer Psychology*, 34(3), 481-491.
- Eyssel, F., & Hegel, F. (2012). (S)he's got the look: Gender stereotyping of robots. *Journal of Applied Social Psychology* 42(9), 2213–2230.
- Fiske, S.T., Cuddy, A.J.C. & Glick, P. (2007). Universal dimensions of social cognition: warmth and competence. *Trends in Cognitive Sciences* 11(2), 77–83.
- 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), e7785.
- Følstad, A., Nordheim, C., & Bjørkli, C. (2018). What makes users trust a chatbot for customer service? An exploratory interview study. In: *The Fifth International Conference on Internet Science – INSCI 2018*, 24 October 2018, 194–208.
- Hentschel, T., Heilman, M.E., & Peus, C.V. (2019). The multiple dimensions of gender stereotypes: A current look at men's and women's characterizations of others and themselves. *Frontiers in Psychology* 10, 11.
- 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.
- Lin, X., Martinengo, L., Jabir, A.I., Ho, A.H.Y., Car, J., Atun, R., & Tudor Car, L. (2023). Scope, Characteristics, behavior change techniques, and quality of conversational agents for mental health and well-being: Systematic assessment of apps. *Journal of Medical Internet Research*, 25, e45984.
- Liu, K., & Tao, D. (2022). The roles of trust, personalization, loss of privacy, and anthropomorphism in public acceptance of smart healthcare services. *Computers in Human Behavior* 127, 107026.
- Nass, C., Steuer, J., & Siminoff, E., (1994). *Computer Are Social Actors*. *Conference on Human Factors in Computing Systems*, CHI 1994, Boston, Massachusetts, USA, April 24-28.
- Park, G., Chung, J., & Lee, S. (2022). Effect of AI chatbot emotional disclosure on user satisfaction and reuse intention for mental health counseling: a serial mediation model. *Current Psychology*, 1-11.
- Pitardi, V., Bartikowski, B., Osburg, V.S., & Yoganathan, V. (2023). Effects of gender congruity in human-robot service interactions: The moderating role of masculinity. *International Journal of Information Management*, 70, 102489.
- Pitardi, V., & Marriott, H.R. (2021). Alexa, she's not human but... Unveiling the drivers of consumers' trust in voice-based artificial intelligence. *Psychology & Marketing*, 38(4), 626–642.
- Schillaci, C.E., De Cosmo, L.M., Piper, L., Nicotra, M., & Guido, G. (2024). Anthropomorphic chatbots' for future healthcare services: Effects of personality, gender, and roles on source credibility, user satisfaction, and intention to use. *Technological Forecasting and Social Change*, 199, 123025.
- Yen, C., & Chiang, M-C. (2021). Trust me, if you can: a study on the factors that influence consumers' purchase intention triggered by chatbots based on brain image evidence and self-reported assessments. *Behaviour & information technology*, 40(11), 1177–1194.
- Zogaj, A., Mähner, P. M., Yang, L., & Tscheulin, D. K. (2023). It's a Match! The effects of chatbot anthropomorphization and chatbot gender on consumer behavior. *Journal of Business Research*, 155, 113412.

## Appendix 1: Stimuli – Study1



**Appendix 2: Measures and their psychometric qualities – Study 1 & 2**

Measure	Study 1		Study 2	
	% variance explained	Chronbach Alpha	% variance explained	Chronbach Alpha
Trust ( <i>Pittardi &amp; Mariott, 2020</i> )	80.4%	0.878	82.31	0.830
Behavioral intentions ( <i>Liu &amp; Tao, 2021</i> )	77.2%	0.850	85.63	0.916
Attitude ( <i>Borau et al., 2021</i> )	92.6%	0.960	95.95	0.979
Competence ( <i>Eyssel &amp; Hegel, 2012</i> )	50.76%	0.750	70.67	0.861
Warmth ( <i>Eyssel &amp; Hegel, 2012</i> )	51,44%	0.843	52.71	0.684
Functional value ( <i>Belanche et al., 2021</i> )	N/A	N/A	75.26	0.934
Emotional value ( <i>Belanche et al., 2021</i> )	N/A	N/A	82.90	0.958