AI or Human? Message Humanness Predicts Perceiving AI as Human: A Secondary Data Analysis of the HeartBot Study

Haruno Suzuki, Jingwen Zhang, Diane Dagyong Kim, Kenji Sagae, Holli A DeVon, Yoshimi Fukuoka

Submitted to: Journal of Medical Internet Research
on: October 29, 2024

Table of Contents

AI or Human? Message Humanness Predicts Perceiving AI as Human: A Secondary Data Analysis of the HeartBot Study

Haruno Suzuki[1] MS; Jingwen Zhang[2] PhD; Diane Dagyong Kim[3] MA; Kenji Sagae[4] PhD; Holli A DeVon[5] PhD; Yoshimi Fukuoka[1] PhD

[1]Department of Physiological Nursing University of California, San Francisco San Francisco US
[2]Department of Communication, Department of Public Health Sciences University of California, Davis Davis US
[3]Department of Communication Davis US
[4]Department of Linguistics University of California, Davis Davis US
[5]University of California, Los Angeles Los Angeles US

Corresponding Author:
Haruno Suzuki MS
Department of Physiological Nursing
University of California, San Francisco
2 Koret Way
San Francisco
US

Abstract

Background: With the rapid advancement of artificial intelligence (AI) technologies in the healthcare field, AI chatbots are advantageous options for improving knowledge and awareness of diseases and modifying health behaviors for diverse populations. However, people's understanding of AI chatbots is still developing, and the factors influencing the perception of AI chatbots and human-AI interaction are largely unknown.

Objective: To identify interaction characteristics related to the perception of believing an AI chatbot as a human versus an artificial agent, controlling for socioeconomic status and past chatbot use in a cohort of diverse women.

Methods: This was a secondary analysis of data from the HeartBot study in women aged 25 years or older. The goal of the HeartBot was to evaluate the change in awareness of heart disease after interacting with a fully automated AI chatbot. Women were recruited through social media from October 2023 to January 2024. The perceived chatbot identity (human vs. artificial argent), length of the HeartBot conversation, humanness in chatbot messages, perception of chatbot message effectiveness, and attitude toward AI were measured at the post-chatbot survey. Multivariate logistic regression was conducted to explore the factors predicting women's perception of a chatbot's identity as a human, adjusting for age, race/ethnicity, education, past chatbot use, humanness in chatbot messages, effectiveness of chatbot messages, and attitude towards AI.

Results: A total of 92 women with a mean age of 45.9 (SD: 11.9, range 26-70) years were analyzed. The chatbot identity was correctly identified by two-thirds (66.3%) of the sample, while one-third (33.7%) incorrectly identified the chatbot as a human. Over half (57.6%) reported having past experiences using a chatbot. Participants interacted with the HeartBot for 13.0 (SD: 7.8) minutes and typed in 82.5 (SD: 61.9) words on average. In the adjusted model, only the score of humanness in chatbot messages was significantly associated with the perception of chatbot identity as a human compared to an artificial argent (adjusted odds ratio 2.37; 95% CI 1.26-4.48; P=.007) controlling for potential confounding factors.

Conclusions: Our findings suggest as chatbot conversations become increasingly natural and humanlike, clearly communicating the chatbot identity to participants is key to establishing correct perceptions. This study offers valuable theoretical and practical insights for the design of AI chatbots in healthcare, emphasizing the important role of message humanness in influencing human perceptions. Future research is warranted to facilitate an understanding of the relationship between chatbot identity, humanness, and health outcomes.

(JMIR Preprints 29/10/2024:67717)
DOI: https://doi.org/10.2196/preprints.67717

Preprint Settings

1) Would you like to publish your submitted manuscript as preprint?

✔ **Please make my preprint PDF available to anyone at any time (recommended).**
Please make my preprint PDF available only to logged-in users; I understand that my title and abstract will remain visible to all users.
Only make the preprint title and abstract visible.
No, I do not wish to publish my submitted manuscript as a preprint.

2) If accepted for publication in a JMIR journal, would you like the PDF to be visible to the public?

✔ **Yes, please make my accepted manuscript PDF available to anyone at any time (Recommended).**
Yes, but please make my accepted manuscript PDF available only to logged-in users; I understand that the title and abstract will remain v
Yes, but only make the title and abstract visible (see Important note, above). I understand that if I later pay to participate in <a href="http

Original Manuscript

Title: AI or Human? Message Humanness Predicts Perceiving AI as Human: A Secondary Data Analysis of the HeartBot Study

Abstract (399/450 words)

Introduction: With the rapid advancement of artificial intelligence (AI) technologies in the healthcare field, AI chatbots are advantageous options for improving knowledge and awareness of diseases and modifying health behaviors for diverse populations. However, people's understanding of AI chatbots is still developing, and the factors influencing the perception of AI chatbots and human-AI interaction are largely unknown.

Objectives: To identify interaction characteristics related to the perception of believing an AI chatbot as a human versus an artificial agent, controlling for socioeconomic status and past chatbot use in a cohort of diverse women.

Methods: This was a secondary analysis of data from the HeartBot study in women aged 25 years or older. The goal of the HeartBot was to evaluate the change in awareness of heart disease after interacting with a fully automated AI chatbot. Women were recruited through social media from October 2023 to January 2024. The perceived chatbot identity (human vs. artificial argent), length of the HeartBot conversation, humanness in chatbot messages, perception of chatbot message effectiveness, and attitude toward AI were measured at the post-chatbot survey. Multivariate logistic regression was conducted to explore the factors predicting women's perception of a chatbot's identity as a human, adjusting for age, race/ethnicity, education, past chatbot use, humanness in chatbot messages, effectiveness of chatbot messages, and attitude towards AI.

Results: A total of 92 women with a mean age of 45.9 (SD: 11.9, range 26-70) years were analyzed. The chatbot identity was correctly identified by two-thirds (66.3%) of the sample, while one-third (33.7%) incorrectly identified the chatbot as a human. Over half (57.6%) reported having past experiences using a chatbot. Participants interacted with the HeartBot for 13.0 (SD: 7.8) minutes and

typed in 82.5 (SD: 61.9) words on average. In the adjusted model, only the score of humanness in chatbot messages was significantly associated with the perception of chatbot identity as a human compared to an artificial argent (adjusted odds ratio 2.37; 95% CI 1.26-4.48; *P*=.007) controlling for potential confounding factors.

Conclusions: Our findings suggest as chatbot conversations become increasingly natural and humanlike, clearly communicating the chatbot identity to participants is key to establishing correct perceptions. This study offers valuable theoretical and practical insights for the design of AI chatbots in healthcare, emphasizing the important role of message humanness in influencing human perceptions. Future research is warranted to facilitate an understanding of the relationship between chatbot identity, humanness, and health outcomes.

Keywords: artificial intelligence; chatbot, women; humanness; chatbot identity; anthropomorphism; user experience

Introduction

In recent years, artificial intelligence (AI) technologies have advanced extensively, and healthcare systems have begun integrating these technologies into clinical practice and research. By combining machine learning and natural language processing, researchers can provide personalized health information and automated counseling to patients.[1–5] Among the various AI applications, chatbots have emerged as one of the most widely tested technologies.

AI chatbots are conversational agents that imitate human interaction through written, oral, and visual communication with users.[6] AI chatbots offer several advantages in increasing health knowledge and modifying health behaviors through natural conversations.[1–5,7–10] These advantages encompass cost-effectiveness, flexible availability, and the ability to reach out to diverse populations.[1] Furthermore, AI chatbots can provide personalized support for health behavior change, including goal setting, continuous monitoring, advice and counseling, sending real-time feedback, and providing on-demand support based on user preferences and behavioral performance data.[7] Clinical trials using AI chatbots have demonstrated the potential efficacy in alleviating depression and anxiety,[1,8,10] reducing the risk of eating disorders,[8] facilitating smoking cessation and medication adherence,[5,7] promoting physical activity and healthy diet,[3,4,7,9] and increasing self-care behaviors.[8,9]

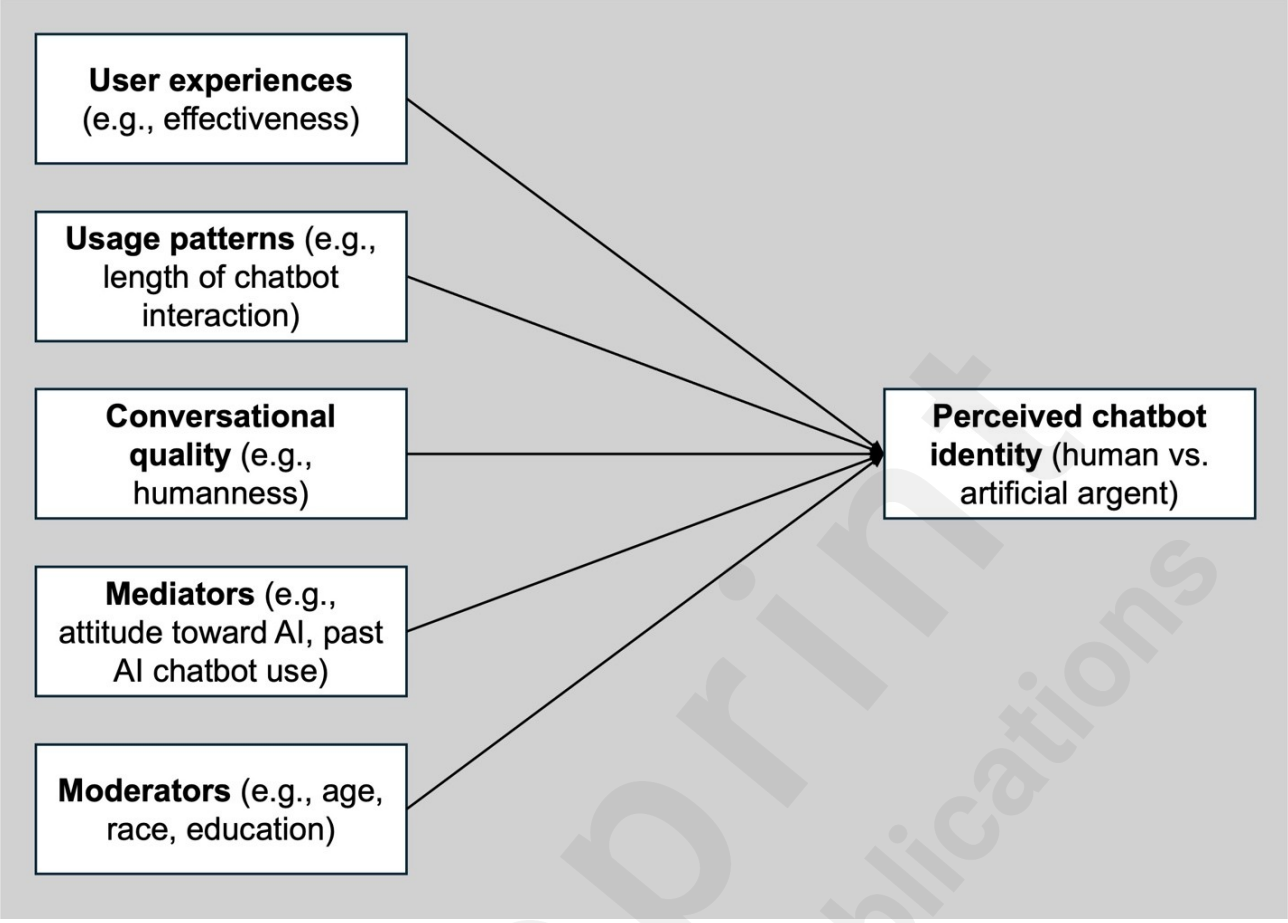
Despite the accumulating evidence showing the promise of AI chatbot health interventions,[1–5,7–10] there are several knowledge gaps about how the quality of human-AI interactions may be contingent upon human perceptions toward the AI chatbot and how actual conversation qualities influence the acceptability, usability, and effectiveness of these interventions. First, few researchers have examined user background factors that may motivate or discourage users' active engagement with an AI chatbot. One study, in a business context, showed users with more positive experiences using general AI tools in the past had a higher tendency to engage more with AI chatbots.[11] Second, limited evidence is available to guide the design of chatbot identity in facilitating chatbot engagement and influence. One recent research study suggests that regardless of the actual message

quality, stereotypical perceptions of artificial agents lacking the capacity to accomplish goals could decrease the effects of AI chatbot intervention messages when the chatbot is labeled as an artificial agent.[12] Another meta-analysis reported that computer-controlled agents were perceived to be less influential than human-controlled avatars in virtual environments, due to perceived less agency from computer programs.[13] These suggest that the effectiveness of chatbot messaging is largely contingent upon the attribution of human agency in chatbot systems.

Relatedly, anthropomorphism, which is the tendency to attribute humanlike characteristics and traits to nonhuman objects, including products, creatures, and natural phenomena,[14] can be an important factor facilitating how humans actively interact with AI chatbots.[15–19] Perceived humanness in AI chatbot content can trigger people to anthropomorphize artificial agents, which may enhance their likability, trustworthiness, and downstream influence effectiveness.[19] In the same vein, in robotics and android studies, the humanlike appearance and movement of robots have been consistently demonstrated to induce positive human empathy. [20,21]

Building upon prior work, the AI chatbot behavior change model[22] provides a valuable framework to guide the design and evaluation of chatbots for health behavior change. It consists of four major components: (1) designing the chatbot characteristics and understanding user backgrounds; (2) building relational capacity; (3) building persuasive conversational capacity; and (4) evaluating mechanisms and outcomes. Guided by this framework, we examined the relationship between perceived AI chatbot identities (human versus artificial agent) and user experiences (e.g., effectiveness, usefulness, satisfaction), usage patterns (e.g., length of chatbot interaction), and conversational quality (e.g., humanness and naturalness of conversation) **[Figure 1]**. We conducted a secondary analysis based on data collected from the HeartBot (a fully automated chatbot) trial, aiming to increase women's heart disease awareness and knowledge of heart attack in community-dwelling women.

Figure 1. Conceptual framework for perceived chatbot identity



Footnote. This framework is adapted from the AI chatbot behavior change model
Abbreviation: AI: artificial intelligence

The HeartBot trial demonstrated the feasibility and acceptability of the program. The main study findings are under review elsewhere. In this trial, the participants were blinded to the chatbot identity (human or AI chatbot) when interacting with a fully-automated HeartBot. This secondary data analysis aimed to examine how women perceive chatbot identity as a human versus a chatbot and to explore the factors associated with different perceptions of chatbot identity in women.

Methods

Design and Sample

We conducted a secondary analysis of the HeartBot trial. The detailed study design and sample eligibility criteria are under review elsewhere. The Institutional Review Board of the sponsoring institution approved the study (No. [Anonymized for Review]) and written informed

consent was obtained from all participants prior to the study. The eligibility criteria included women aged 25 years or older, residing in the United States, proficient in English, possessing a cell phone with texting capabilities, having internet access, without self-reported cognitive impairment or a history of heart disease or stroke, and not being a healthcare provider or student in a healthcare related field. We followed STROBE reporting guidelines.[23]

Procedure

Participants were recruited through social media (i.e., Facebook and Instagram) ads placed from October 2023 to January 2024, using targeting strategies that aim to reach racially and ethnically diverse demographics (e.g., Hispanic and Black/African American women). Those interested in the study were redirected to our study site, where they reviewed and the study aims and procedures. Those who agreed to be contacted by the research team completed an online eligibility survey. An e-consent form was sent to those who met the eligibility criteria. After obtaining the e-consent, participants were asked to complete an online baseline survey consisting of sociodemographic status, cardiovascular risk factors, medication intake, and past AI chatbot use. Participants were then instructed to initiate a text-messaging-based conversation with HeartBot, available 24 hours a day, 7 days a week. Research staff monitored the conversations between HeartBot and participants to ensure participants' safety by verifying that the information provided by HeartBot was accurate. After 4 to 6 week of the HeartBot interaction, participants were asked to complete an online post-intervention survey, including AI chatbot interaction and evaluations. All online surveys were administered by Research Electronic Data Capture (REDCap),[24] a secure online tool utilized to manage study data. Upon completion of all study procedures, participants received a $20 e-gift card.

Measures

Baseline Measures

Sociodemographic, cardiovascular risks, medication, and past AI chatbot use

Sociodemographic factors, such as age, race/ethnicity, education, household income, marital status, employment status, and immigration experience to the United States, were collected from participants in the baseline survey.

Data collected at baseline included; self-reported cardiovascular risks, including menopause, Body Mass Index (BMI, kg/m^2), calculated with height and weight, smoking in the past 30 days; physical activity ≥150 mins per week; family history of heart disease; prescribed blood pressure; cholesterol; diabetes medication; and daily aspirin intake. The cardiovascular risk factor variables were selected based on the latest clinical guidelines.[25] We assessed past AI chatbot use experience with the following question "*Have you used any chatbot in the past 30 days*?" There were 2 response options: 1) Yes, 2) No.

Post-intervention Measures

AI chatbot interaction

In the post-survey, we measured several metrics denoting users' interaction patterns with an AI chatbot, including users' word count, time in minutes spent in the conversation, and the number of questions asked by users.

Humanness in chatbot messages

In the AI chatbot behavior change model,[22] humanness in chatbot messages is categorized as the "conversational quality," which measures the degree of perceived humanness in a chatbot conversation. Participants rated the humanness in chatbot messages using the "Anthropomorphism Scale"[21] in the post-intervention survey. The scale consists of 5 items (natural vs. fake, humanlike vs.

machine-like, conscious vs. unconscious, lifelike vs. artificial, adaptive vs. rigid) using a 7-point Likert scale based on a horizontal visual analog scale.[21] The scores on the scale were summed and averaged to create a mean composite score. A higher score indicates more humanlike HeartBot messages. The scale was developed based on a previous study.[21] The internal consistency of the scale was strong with Cronbach's alpha = 0.90 in our study sample.

Effectiveness of chatbot messages

In the AI chatbot behavior change model,[22] effectiveness of chatbot messages is categorized as the "user experiences", which measure the level of usefulness and convenience in chatbot conversations. Participants rated the self-reported effectiveness of chatbot messages using the "Effectiveness Scale" in the post-intervention survey. The scale was originally developed based on previous literature.[12,26] The scale consists of 5 items (effective vs. ineffective, helpful vs. unhelpful, beneficial vs. not beneficial, adequate vs. not adequate, supportive vs. not supportive) using a 7-point Likert scale based on a horizontal visual analog scale.[12,26] The scores on the scale were summed and averaged to create a mean composite score. A higher score indicates greater effectiveness of the HeartBot messages. The internal consistency of the scale was strong as evidenced by Cronbach's alpha = 0.93 in our study sample.

Attitude toward AI

To investigate the attitude toward AI chatbots, participants were asked the following question on the post-intervention survey: "*How positive or negative do you feel about the use of artificial intelligence in healthcare?*" There are 5 response options: 1) very positive, 2) positive, 3) neutral, 4) negative, 5) very negative.

Perception of chatbot identity (human versus AI chatbot)

To determine the perception of the identity of HeartBot, participants were asked the following question at the post-intervention survey: *"Do you think you texted a human or an artificial intelligent chatbot during your conversation?"* There were 2 response options: 1) human, 2) artificial

agent.

Statistical Analysis

Descriptive analyses were used to describe participants' sociodemographic backgrounds, cardiovascular risks, medication, and AI chatbot interactions and evaluations. The study sample was split based on the perception of chatbot identity as a human versus an artificial argent. Chi-squared, Fisher's exact test, and Wilcoxson rank-sum tests were used to compare the differences in baseline sample characteristics of the two subsamples.

Race/ethnicity and education were recoded into dichotomous variables: non-White or White and "completed college/graduate school" or "less than high school/didn't complete college," respectively, in a logistic regression analysis. Attitude toward AI chatbot was divided into three categories: positive, neutral, and negative. Recoding several variables was aimed at improving statistical power. Additionally, univariate logistic regression analyses were performed to estimate the relationships between the dependent variable (perception of chatbot identity) and each independent variable with sociodemographic factors, cardiovascular risk factors, medication intake, AI chatbot interaction, and AI chatbot evaluation.

Finally, a multivariate logistic regression analysis was conducted to explore the factors that were significantly associated with participants' perception of chatbot identity as being a human. The independent variables ensured face validity (e.g., age, race, education), and the potential confounders referred to in literature[25] were entered into a multivariate regression model. The potential confounders to predict the perception of chatbot identity were selected based on the AI chatbot behavior change model,[22] including past chatbot use, conversation length of chatbots, humanness in chatbot messages, perceived effectiveness of messages, and attitude toward AI. All predictor variables were included in this regression model to control for the effects of each potential covariate. The rationale for selecting potential covariates in the regression model was based on previous literature.[25] Multicollinearity was tested to ensure that independent variables were not highly

correlated. The variance inflation factor values of all independent variables ranged from 1.13 to 2.12 (mean:1.47), indicating an acceptable range and no multicollinearity in the variables. Statistical significance was set at a two-sided p-value < 0.05. All analyses were performed using Stata v18.0.[27]

Results

Sample Characteristics

A total of 92 participants completed the baseline, HeartBot interaction, and post-intervention surveys **[Table 1]**. The mean age (Standard Deviation [SD], range) of participants was 45.9 (11.9, 26-70) years. 40.2% of participants identified their race/ethnicity as White/Caucasian, 23.9% as Black/African American, and 20.7% as Hispanic/Latino Americans. 71.7% reported completing college or graduate school. 44.6% of participants reported experiencing menopause. 36.3% of participants reported their BMI was 30 or above. 27.2% reported taking blood pressure medication. 57.6% reported experiencing a prior interaction with an AI chatbot. The most popular types of chatbots were ChatGPT (23.9%) and Siri (21.7%).

Table 1. Sample characteristics in respondent perception of chatbot identity as a human vs. an artificial agent (N=92)

	Mean (SD) [range] or n (%)			
	Overall (N=92)	Human (n=31)	Artificial agent (n=61)	*P*-value[a]
Sociodemographic factors				
Age (year)	45.9 (11.9) [26-70]	46.3 (12.2) [28-70]	45.6 (11.9) [26-68]	.816
Race/ethnicity				
American Indian/Alaskan Native	1 (1.1)	0 (0.0)	1 (1.6)	.410
Asian	6 (6.5)	4 (12.9)	2 (3.3)	
Black/African American	22 (23.9)	9 (29.0)	13 (21.3)	
Hispanic/Latino	19 (20.7)	4 (12.9)	15 (24.6)	
Native Hawaiian/Other Pacific Islander	2 (2.2)	0 (0.0)	2 (3.3)	
White/Caucasian	37 (40.2)	12 (38.7)	25 (41.0)	
More than 1 race/ethnicity	5 (5.4)	2 (6.5)	3 (4.9)	
Education				
No more than high school/didn't complete college	26 (28.3)	11 (35.5)	15 (24.6)	.273
Completed college/graduate school	66 (71.7)	20 (64.5)	46 (75.4)	
Household income				
Less than $75,000/ Don't know / decline to respond	33 (57.6)	21 (67.7)	32 (52.5)	.161
$75,000 or above	39 (42.4)	10 (32.3)	29 (47.5)	
Marital status				

Never married	21 (22.8)	8 (25.8)	13 (21.3)	.764
Currently married/cohabitating	59 (64.1)	20 (64.5)	39 (63.9)	
Divorced/widowed	12 (13.0)	3 (9.7)	9 (14.8)	
Employment status				
Employed for full time/part time	56 (60.9)	18 (58.1)	38 (62.3)	.769
Unemployed/Looking for a job/Student/Homemaker	17 (18.5)	7 (22.6)	10 (16.4)	
Retired/Disabled/Other	19 (20.7)	6 (19.4)	13 (21.3)	
Immigration experience to the United States	12 (13.0)	5 (16.1)	7 (11.5)	.374
Cardiovascular risk factors / Medication intake				
Menopause	41 (44.6)	14 (45.2)	27 (44.3)	.935
Body Mass Index (kg/m^2)				
Less than 30	58 (63.7)	16 (51.6)	42 (70.0)	.084
30 or above	33 (36.3)	15 (48.4)	18 (30.0)	
Smoking in the past 30 days	14 (15.2)	4 (12.9)	10 (16.4)	.457
Physical activity ≥150 mins per week	56 (60.9)	20 (64.5)	36 (59.0)	.609
Family history of heart disease	13 (14.1)	4 (12.9)	9 (14.8)	.540
Blood pressure medication	25 (27.2)	6 (19.4)	19 (31.2)	.229
Cholesterol medication	16 (17.4)	5 (16.1)	11 (18.0)	.820
Diabetes medication	17 (18.5)	6 (19.4)	11 (18.0)	.877
Taking aspirin daily	13 (14.1)	5 (16.1)	8 (13.1)	.460
AI chatbot interaction				
Past AI chatbot use	53 (57.6)	17 (54.8)	36 (59.0)	.702
Conversation length (words)	82.5 (61.9) [34-377]	81.8 (67.0) [36-360]	82.8 (59.8) [34-377]	.183
Conversation length (minute)	13.0 (7.8) [5.6-42.2]	13.1 (9.6) [5.6-42.2]	12.9 (6.8) [5.6-40.3]	.329
Number of questions asked to HeartBot (at least one)	27 (29.4)	7 (22.6)	20 (32.8)	.310
AI chatbot evaluation				
Humanness in chatbot message	5.2 (1.2) [2.0-7.0]	5.7 (1.1) [3.4-7.0]	4.9 (1.2) [2.0-7.0]	.003
Effectiveness of chatbot messages	5.7 (1.2) [1.0-7.0]	5.9 (0.9) [3.4-7.0]	5.6 (1.4) [1.0-7.0]	.619
Attitude toward artificial intelligence				
Positive	35 (38.0)	12 (38.7)	23 (37.7)	1.000
Neutral	44 (47.8)	15 (48.4)	29 (47.5)	
Negative	13 (14.1)	4 (12.9)	9 (14.8)	

[a]*P*-value was calculated by chi-squared test, Fisher's exact test, or Wilcoxon rank-sum test.

HeartBot Interaction

While 33.7 % identified the chatbot as a human, 66.3% reported they interacted with an artificial agent. The mean (SD, range) and median (Interquartile Range) of conversation length with HeartBot by word count and minute were 82.5 (61.9, 34-377), 64.5 (48.0) words and 13.0 (7.8, 5.6-42.2), 10.6 (5.4) minutes. The mean (SD, range) scores of humanness in chatbot messages and effectiveness of chatbot messages were 5.2 (1.2, 2.0-7.0) and 5.7 (1.2, 1.0-7.0), respectively. 38.0%

had a positive feeling for AI. In the bivariate analysis, the mean score of humanness in chatbot messages was significantly higher in the group who answered the chatbot identity as a human compared to the group who thought they were interacting with an artificial agent (*P*=.003).

Table 2 presents the unadjusted and adjusted odds ratios from multivariate logistic regression analysis results for predicting the perception of chatbot identity as a human versus a chatbot. In the unadjusted model, the score of humanness in chatbot messages was significantly associated with the perception of chatbot identity as a human compared to an artificial argent (unadjusted odds ratio 1.81; 95% confidence interval [CI] 1.19-2.77; *P*=.006). In the adjusted model, only the score of humanness in chatbot messages was significantly associated with the perception of chatbot identity as a human compared to an artificial argent (adjusted odds ratio 2.37; 95% CI 1.26-4.48; *P*=.007), controlling for age, race/ethnicity, education, past chatbot use, conversation length in a chatbot, the effectiveness of the message, and attitude toward AI.

Table 2. Unadjusted and adjusted odds ratios (ORs) from multivariate logistic regression analysis for predicting the perception of chatbot identity as being a human (N=92)

Variables	OR (95% CI)	*P* value	AOR (95% CI)	*P* value
Age	1.01 (.97, 1.04)	.785	.99 (.95, 1.04)	.796
Race/ethnicity				
White	1 [Reference]		1 [Reference]	
Non-White[a]	1.10 (.45, 2.66)	.833	1.15 (.37, 3.57)	.809
Education				
Less than high school/didn't complete college	1 [Reference]		1 [Reference]	
Completed college/graduate school	.59 (.23, 1.52)	.275	.56 (.19, 1.66)	.292
Past AI chatbot use				
No	1 [Reference]		1 [Reference]	
Yes	.84 (.35, 2.02)	.702	.93 (.31, 2.79)	.901
Conversation length (words)	1.00 (.99, 1.01)	.940	1.00 (.99, 1.01)	.930
Humanness in chatbot messages	1.81 (1.19, 2.77)	.006*	2.37 (1.26, 4.48)	.007*
Effectiveness of chatbot messages	1.23 (.84, 1.81)	.289	.70 (.37, 1.33)	.279
Attitude toward artificial intelligence				
Negative	1 [Reference]		1 [Reference]	
Neutral	1.16 (.31, 4.41)	.823	1.16 (.22, 6.07)	.865
Positive	1.17 (.30, 4.62)	.818	1.01 (.16, 6.43)	.989

[a]Non-White included American Indian, Alaskan Native, Asian, Black/African American,

Hispanic/Latino, Native Hawaiian, other Pacific Islander, and multiracial individuals.

Discussion

Principal findings

This study explored whether and why people attribute different identities to an AI chatbot, specifically the extent to which AI chatbots are perceived as human versus artificial agents among women, and identified key factors influencing these perceptions. A key finding in this study was that individuals who perceived a higher degree of humanness in chatbot messages were more likely to identify the chatbot as human. In contrast, neither the perceived effectiveness of the chatbot message nor general attitudes toward AI influenced the perception of the chatbot identity. These results suggest that the naturalness and humanness of chatbot communication play a critical role in shaping users' perceptions of chatbot identity. To the best of our knowledge, this is the first study to examine the relationship between perceived humanness in chatbot messages and chatbot identity using data from a fully automated AI chatbot within a healthcare context, contributing novel insights to the field of human-computer interactions in healthcare.

The current study demonstrated that perceived humanness in chatbot messages is associated with the user's tendency to perceive chatbot as being a human. This finding is consistent with existing research findings. According to Go and Sundar,[28] there are three factors that influence humanness among AI chatbots: visual cues (e.g., using human figures), conversational cues (e.g., interactive or contingent messages), and identity cues (e.g., humanlike names or identities). Conversational cues refer to humanlike contingent conversational markers that increase expectations for humanlike communication with chatbots, such as using empathetic phrases, polite statements, and acknowledging users' previous responses.[29] Such conversational cues can enhance the naturalness and humanness of the conversations delivered by chatbots. Our findings provide empirical evidence that perceived humanness through conversations does play a crucial role in determining a user's perception of the chatbot's identity.

This is important because how users perceive a chatbot's identity can have implications on their expectations and evaluations of the chatbot's performance and effectiveness. If users assume the chatbot identity as an artificial agent, they are more likely to assess the quality of chatbot performance based on their existing stereotypes of chatbots.[28] In contrast, if they assume a chatbot is human, they are more likely to assess the quality of chatbot performance based on their expectations of other humans. Usually, when users perceive a chatbot more as a human, they expect better and more natural performance from the chatbot than when they perceive the chatbot as an AI agent.[30]

Subjective expectancy of the chatbot's performance matters, and if expectancy is not met, user evaluations of the chatbot will be poorer. This is explained by the "expectancy violation effect."[31,32] Thus, calibrating users' expectations of AI chatbots is an important consideration in designing chatbots' characteristics and conversational features. For instance, past research has shown that message contingency, defined as humanlike continuous dialogues remembering previous responses, could enhance a chatbot's social presence, which further increased users' perceived intelligence and friendliness of the chatbot.[28] In our study, even though all participants were informed that they were interacting with a chatbot named "Heartbot," a significant portion of the women still thought they were interacting with a human. This "misperception" could be due to the fact that they evaluated the chatbot messages' naturalness and humanness highly.

The effectiveness of chatbot messages and attitude towards AI were not associated with the perception of chatbot identity in this study. A previous study similarly revealed that the actual performance level of a chatbot did not influence perceiving a chatbot as a human.[33] This suggests that perceiving an AI chatbot as a human or a computer agent is largely dependent upon the encoding and decoding of the actual conversational messages, rather than the conversational context or the impact of the conversation.

Considering AI chatbot applications in the healthcare field, the current study emphasizes the necessity to clearly disclose and inform the users of the chatbot identity, especially given that AI's

conversational capacity is only going to improve in the future. Although earlier research showed using human identity or strategically hiding AI chatbot's identity may be advantageous in enhancing the user experience,[28,29,34–36] we argue it is no longer ethical and can be counterproductive in forming the right expectations and useful interactions with AI chatbots.

The current study findings provide evidence to support a theoretical basis for explaining human-AI interaction and to design effective interventions using an AI chatbot to improve health behaviors and outcomes, from health education to counseling support in e-health. Future studies are needed to evaluate both conversation quality and the acceptability and efficiency of health AI chatbot intervention. In addition, further studies are warranted to support the present findings that examine the relationship between the humanness of chatbot messages and chatbot identity.

Strengths and Limitations

The strengths of this study are that we collected detailed data on user experiences and perceptions of AI chatbots based on a rigorous theoretical framework and recruited diverse racial and ethnic women in the U.S. However, several limitations should be acknowledged. First, the convenience sampling method may have affected the study findings by selection bias. Second, only female adults in the U.S. were included in the HeartBot program. The current findings may not be generalizable to male adults and other country's populations. Lastly, the humanness of chatbot message and perceived effectiveness of chatbot message are self-report instruments. Therefore, it may not reflect the objective condition of using the HeartBot. Future studies are needed to combine subjective and objective data to evaluate the acceptability and usability of AI chatbot interventions.

Conclusion

This study highlights the significant role that perceived humanness in chatbot messages plays in shaping the user's perception of chatbot identity. Conversely, the effectiveness of chatbot messages and attitudes toward AI did not significantly influence the perception of the chatbot identity as a human. This suggests that the perceived humanlike attributes primarily drive users to

attribute a human identity to the chatbot, specifically in healthcare settings where user trust and engagement are crucial. The current study provides a theoretical foundation for understanding human-AI interactions and offers practical insights for designing user-friendly AI chatbots in healthcare. Further research is needed to explore the relationship between chatbot humanness, identity, and user outcomes to optimize the design and implementation of AI chatbots in health-related fields.

Acknowledgments

The study was supported by the Noyce Foundation and the UCSF School of Nursing GAIN Fund. The study sponsors had no role in the study design, data collection, data analysis, manuscript writing, or decision to submit the report for publication.

HS, JZ, DDK, KS, HAD, and YF contributed to the conception and design and wrote sections of the manuscript. DDK and YF contributed to collecting all data. HS and YF contributed to statistical analysis. All authors reviewed the manuscript and approved the submitted version.

Conflicts of Interest

None declared.

Abbreviations

AI: artificial intelligence

BMI: Body Mass Index

CI: confidence interval

SD: standard deviation

References

1. Zhong W, Luo J, Zhang H. The therapeutic effectiveness of artificial intelligence-based chatbots in alleviation of depressive and anxiety symptoms in short-course treatments: A systematic review and meta-analysis. Journal of Affective Disorders. 2024;356:459-469. doi:10.1016/j.jad.2024.04.057

2. Kurniawan MH, Handiyani H, Nuraini T, Hariyati RTS, Sutrisno S. A systematic review of artificial intelligence-powered (AI-powered) chatbot intervention for managing chronic illness. Annals of Medicine. 2024;56(1):2302980. doi:10.1080/07853890.2024.2302980

3. Oh YJ, Zhang J, Fang ML, Fukuoka Y. A systematic review of artificial intelligence chatbots for promoting physical activity, healthy diet, and weight loss. International Journal of Behavioral Nutrition and Physical Activity. 2021;18(1):160. doi:10.1186/s12966-021-01224-6

4. Noh E, Won J, Jo S, Hahm DH, Lee H. Conversational agents for body weight management: systematic review. J Med Internet Res. 2023;25(1):e42238. doi:10.2196/42238

5. Bendotti H, Lawler S, Chan GCK, Gartner C, Ireland D, Marshall HM. Conversational artificial intelligence interventions to support smoking cessation: A systematic review and meta-analysis. Digit Health. 2023;9:20552076231211634. doi:10.1177/20552076231211634

6. Laranjo L, Dunn AG, Tong HL, et al. Conversational agents in healthcare: a systematic review. J Am Med Inform Assoc. 2018;25(9):1248-1258. doi:10.1093/jamia/ocy072

7. Aggarwal A, Tam CC, Wu D, Li X, Qiao S. Artificial intelligence-based chatbots for promoting health behavioral changes: systematic review. J Med Internet Res. 2023;25:e40789. doi:10.2196/40789

8. Kim HK. The effects of artificial intelligence chatbots on women's health: a systematic review and meta-analysis. Healthcare (Basel). 2024;12(5):534. doi:10.3390/healthcare12050534

9. Lyzwinski LN, Elgendi M, Menon C. Conversational agents and avatars for cardiometabolic risk factors and lifestyle-related behaviors: scoping review. JMIR Mhealth Uhealth. 2023;11:e39649. doi:10.2196/39649

10. He Y, Yang L, Qian C, et al. Conversational agent interventions for mental health problems: systematic review and meta-analysis of randomized controlled trials. J Med Internet Res. 2023;25:e43862. doi:10.2196/43862

11. Kasilingam DL. Understanding the attitude and intention to use smartphone chatbots for shopping. Technology in Society. 2020;62:101280. doi:10.1016/j.techsoc.2020.101280

12. Liao W, Oh YJ, Feng B, Zhang J. Understanding the influence discrepancy between human and artificial agent in advice interactions: the role of stereotypical perception of agency. Communication Research. 2023;50(5):633-664. doi:10.1177/00936502221138427

13. Fox J, Ahn SJ (Grace), Janssen JH, Yeykelis L, Segovia KY, Bailenson JN. Avatars versus agents: a meta-analysis quantifying the effect of agency on social influence. Human–Computer Interaction. 2015;30(5):401-432. doi:10.1080/07370024.2014.921494

14. Epley N, Waytz A, Cacioppo JT. On seeing human: a three-factor theory of anthropomorphism. Psychol Rev. 2007;114(4):864-886. doi:10.1037/0033-295X.114.4.864

15. Han MC. The Impact of Anthropomorphism on consumers' purchase decision in chatbot commerce. Journal of Internet Commerce. 2021;20(1):46-65. doi:10.1080/15332861.2020.1863022

16. Klein K, Martinez LF. The impact of anthropomorphism on customer satisfaction in chatbot commerce: an experimental study in the food sector. Electron Commer Res. 2023;23(4):2789-2825. doi:10.1007/s10660-022-09562-8

17. Roy R, Naidoo V. Enhancing chatbot effectiveness: The role of anthropomorphic

conversational styles and time orientation. Journal of Business Research. 2021;126:23-34. doi:10.1016/j.jbusres.2020.12.051

18. Konya-Baumbach E, Biller M, von Janda S. Someone out there? A study on the social presence of anthropomorphized chatbots. Comput Hum Behav. 2023;139. doi:10.1016/j.chb.2022.107513

19. Cheng X, Zhang X, Cohen J, Mou J. Human vs. AI: Understanding the impact of anthropomorphism on consumer response to chatbots from the perspective of trust and relationship norms. Information Processing & Management. 2022;59(3):102940. doi:10.1016/j.ipm.2022.102940

20. Ishiguro H. Android science: Toward a new cross-interdisciplinary framework. In: Thrun S, Brooks R, Durrant-Whyte H. Robotics research: results of the 12th International Symposium ISRR. Berlin, Heidelberg, New York: Springer; 2007. p. 118-27. ISBN: 3540481109

21. Bartneck C, Kulić D, Croft E, Zoghbi S. Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots. Int J of Soc Robotics. 2009;1(1):71-81. doi:10.1007/s12369-008-0001-3

22. Zhang J, Oh YJ, Lange P, Yu Z, Fukuoka Y. Artificial intelligence chatbot behavior change model for designing artificial intelligence chatbots to promote physical activity and a healthy diet: Viewpoint. J Med Internet Res. 2020;22(9):e22845. doi:10.2196/22845

23. Elm E von, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. Strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. BMJ. 2007;335(7624):806-808. doi:10.1136/bmj.39335.541782.AD

24. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)—A metadata-driven methodology and workflow process for providing translational research informatics support. Journal of Biomedical Informatics. 2009;42(2):377-381. doi:10.1016/j.jbi.2008.08.010

25. Martin SS, Aday AW, Almarzooq ZI, et al. 2024 Heart disease and stroke statistics: A report of US and global data from the American Heart Association. Circulation. 2024;149(8):e347-e913. doi:10.1161/CIR.0000000000001209

26. Feng B. Testing an integrated model of advice giving in supportive interactions. Human Communication Research. 2009;35(1):115-129. doi:10.1111/j.1468-2958.2008.01340.x

27. SoftwareStataCorpLLC. Stata Statistical Software. 2023.

28. Go E, Sundar SS. Humanizing chatbots: The effects of visual, identity and conversational cues on humanness perceptions. Comput Hum Behav. 2019;97(C):304-316. doi:10.1016/j.chb.2019.01.020

29. Svenningsson N, Faraon M. Artificial intelligence in conversational agents: A study of factors related to perceived humanness in chatbots. Proceedings of the 2019 2nd Artificial Intelligence and Cloud Computing Conference. AICCC '19. Association for Computing Machinery; 2020:151-161. doi:10.1145/3375959.3375973

30. Sundar SS. The MAIN Model: A heuristic approach to understanding technology effects on credibility. Ed: Metzger MJ and Flanagin AJ. MacArthur Foundation Series on Digital Media and Learning. Cambridge, MA: The MIT Press; 2008. p. 73-100. doi: 10.1162/dmal.9780262562324.073

31. Cappella JN, Greene JO. A discrepancy-arousal explanation of mutual influence in expressive behavior for adult and infant-adult interaction 1. Communication Monographs. 1982;49(2):89-114. doi:10.1080/03637758209376074

32. Short J, Williams E, Christie B. The social psychology of telecommunications. London: Wiley; 1976. ISBN: 0471015814

33. Koh YJ, Sundar SS. Heuristic versus systematic processing of specialist versus generalist sources in online media. Human Communication Research. 2010;36(2):103-124. doi:10.1111/j.1468-2958.2010.01370.x

34. Sundar SS, Bellur S, Oh J, Jia H, Kim HS. Theoretical importance of contingency in human-computer interaction: Effects of message interactivity on user engagement. Communication Research. 2016;43(5):595-625. doi:10.1177/0093650214534962

35. Wuenderlich N, Paluch S. A nice and friendly chat with a bot: User perceptions of AI-based service agents. Proceeding of the ICIS 2017; 2017 Dec 10. https://aisel.aisnet.org/icis2017/ServiceScience/Presentations/11 Accessed October 29, 2024.

36. Luo X, Tong S, Fang Z, Qu Z. Machines versus humans: The impact of AI chatbot disclosure on customer purchases. Published online July 1, 2019. doi:10.2139/ssrn.3435635

Supplementary Files

Figures

Conceptual framework for perceived chatbot identity.