Effects of a Chatbot's Appearance and Task-Type on Social Presence, Perceived Accuracy, Credibility, and Likeliness to Use the Chatbot Again

Gianluca Cabria, 437693

Department of Communication and Information Sciences, Tilburg University

Dr. Mincheol Shin

30/08/2022

Abstract

Given the general uncertainty about the effects that chatbots presenting specific characteristics may have on users, this study aimed to investigate the effects of interacting with a human-like chatbot compared to a robotic-like one. A 2 (Appearance: Anthropomorphic vs. Robotic) x 2 (Type of task: Human vs. Mechanical) between-subjects experiment (N = 120) was conducted to examine the effects of the chatbot's appearance (i.e., anthropomorphic) on social presence, perceived credibility, perceived accuracy, and likeliness to use the chatbot again. In addition, the type of task the agent had to perform (either mechanical or human) was scrutinized as a possible moderator between the chatbot's appearance and the users' overall perception. At the same time, the extent of perceived social presence and credibility/accuracy were posited as plausible sequential mediators between the agent's appearance and the likeliness to reuse it. In accordance with previous work, results showed that combining anthropomorphic appearance with a human task boosted the user's likeliness to reuse the chatbot to the greatest extent among the four conditions. The sequential mediation hypothesis, however, found no significant support. Further implications of the findings were discussed along with previous findings in the field.

Keywords: chatbots, appearance, type of task, social presence, perceived credibility, perceived accuracy

Acknowledgements

To my Dad, and my Mum.

To Mincheol, whose immensely kind supervision helped me arriving where I am now.

To Pink Floyd, Jimi Hendrix, & Nietzsche.

To Angelo, beloved mentor.

To Rebeca, who accepted me for who I am.

And to all my friends, from first to last.

Table of contents

1. Introduction

It is no secret that technology is constantly and rapidly advancing. Until a few decades ago, humans were used to having a certain degree of control over technology. This belief is increasingly being questioned and studied by recent scholars, as it is now possible to perform a multitude of actions in a limited period of time thanks to specific interactive agents (e.g., Alexa; Woebot). The increasing reliance on such highly interactive AI, has brought the related field of research and users themselves to question the role of the new agents in everyday life. The relational interdependence between us humans and AI, in fact, seems not limited to the number of tasks that the latter can do. Rather, besides the tasks AI is nowadays able to do, it is *how* it performs. Specifically, users tend to increasingly rely on such technologies to carry out almost any kind of activity ranging from shopping to private personal consultancy (Lee & See, 2004). The factors influencing this interdependence, however, are not completely clear to researchers.

Such a paradigmatic shift in the AI-user relationship is related to the increased autonomy that many entities within the IoT environment have acquired in the past years (Kang & Kim, 2022). More precisely, we have witnessed the evolution of passive electronic tools changing into artificial intelligence entities that can provide highly-reliable information based on algorithmic operations. As algorithms have now reached a point where they can manipulate the information gathered (e.g., suggesting an ideal weekly schedule based on a user's frequent and specific input rather than booking a flight ticket autonomously), users have become more and more likely to perceive digital agents as entities with which it became possible to interact as if they are interacting with humans (Kang & Kim, 2022). Sundar and Nass (2001) reported that this distinct "sourceness" originated from the constant human-agent interaction, which in turn allowed the machine agency to be prominent alongside the human one. Therefore, it is more than an

assumption to state that one of the main elements influencing this constant increasing interdependence between humans and AI is the latter's algorithmic hyper-efficiency (Gal-Ezer & Zur, 2004).

A significant example of interactive agents whose usage has been demonstrated to be in constant increment are *chatbots* (Murtarelli et al., 2022). A chatbot is a computer program capable of responding when "conversed with through text or voice while understanding several languages by Natural Language Processing" (Khanna et al., 2015, p. 2). A more lexical demarcation is the one that considers chatbots as smart bots; digital assistants; interactive agents; and artificial conversational entities designed to stimulate conversations with humans over the Internet (Adamopoulou & Moussiades, 2020).

According to specific chatbot-related studies, there are several dimensions besides algorithmic hyper-efficiency that might have an impact on the users' agent perception. Two of these main dimensions include the agents' appearance, (Glikson & Woolley, 2020) and the type of task they are programmed to execute (Davenport et al., 2020). With appearance, we refer to the superficial properties of chatbots (Fulmer et al., 2018). An agent with human appearance, for instance, would be characterized by *cues* such as avatars presenting human characteristics. On the other hand, an agent with a robotic appearance would be characterized by cues such as avatars with robotic appearance. While defining task type instead, we refer to specific actions an agent performs in relation to the context where it is employed. For instance, an e-commerce task would be considered typically mechanical. At the same time, a cognitive therapy-related one as human, as the latter supposedly requires an extent of coded human behavioral understanding (Bell et al., 2019).

For some time now, researchers have tried to grasp how technological features would affect users' daily lives in the long term, as the rise of the machine agency has disrupted our daily habits (Kang & Kim, 2022). In the field, a large number of sectors, including healthcare (Nadarzynski et al., 2019), finance (Ortiz, 2021), government-citizen communication (Androutsopoulou et al., 2019), and engineering (Telang et al., 2018), is increasingly relying on chatbot technologies. However, although several studies have proven several contrasting effects of chatbots' appearance in determined contexts such as finance and engineering (Kang & Kim, 2022; Han, 2021; Morana et al., 2020; Fulmer et al., 2018), researchers tend to analyze situations in which chatbots perform mechanical tasks rather than human tasks (Khan & Rabbani, 2021; Cho & Yun, 2019; Chichanowsky et al., 2018; Brandtzaeg & Følstad, 2017). As a consequence, results concerning the effectiveness and users' preferences within the automated human task field are still recent and mixed (Oh et al., 2020).

Accordingly, one of the areas where research is needed is the comparison between the effects of the outcomes of chatbots performing cognitive behavioral (CB) psychology-related tasks with those performing mechanical tasks on the users' preferences. Although purportedly scientific, psychology requires a certain degree of humanity during, for example, a therapeutic consultancy. In fact, it is commonly believed that chatbots are best suited for executing tasks that do not require emotional thinking, as they tend to be perceived as "unemotional and cold" (Sundar, 2020). Nonetheless, although CBT-related tasks can be considered harder than mechanical ones, the topic of self-disclosure preferences should always be mentioned. Extensive research found that some people, while being aware that AI therapeutical consultancy might not be as accurate as that provided by a human practitioner, would rather talk to an agent to avoid embarrassing situations (Lee et al., 2020; Van der Lee et al., 2019).

Therefore, depending on the users' context, the outcomes in this research field yielded mixed results (Fulmer et al., 2018; Liu et al., 2022; Mathur et al., 2022). In this respect, Fulmer et al. (2018) found that chatbots involved in more human tasks can trigger mixed reactions concerning engagement and user satisfaction based on their appearance. Mixed results were also yielded by Nadarzinski's et al. (2019) experiment, which underlined significant individual differences in the understanding of chatbots, AI hesitancy, and motivation to use health chatbots in comparison to agents performing more mechanical than human tasks.

As mentioned already, multiple factors that may influence the user's AI perception. Specifically, elements such as the chatbot's appearance and type of task tend to interact, although the results of such interaction vary based on the context in which the agent operates (Nadarzinski's et al. 2019; Fulmer et al., 2018). Given the increasing reliance on chatbots and the gap in the literature considering the contrast between mechanical-human tasks and appearance interaction, the comparison between CBT and "mechanical"-related chatbots lacks sufficient empirical evidence while remaining understudied. Taken together, this paper will investigate whether the interaction between a chatbot's type of task (mechanical vs.. human) and its appearance (robotic vs.. human) may induce any significant effects on user perceptions drawing on the concept of machine heuristics that will be extensively explained in the next section.

1. Literature Review and Hypotheses

In an attempt to provide valid explanations for specific users' attributions to chatbots, researchers not only focused on the dimensions -such as their appearance or type of task- strengthening this relationship but also tried instead to break down the two-sided dependency into smaller psychological dimensions. For example, Sundar (2008) developed a model that predicts the potential effects of the structural features in new interactive agents on user

perceptions and behaviors, particularly focusing on the interrelation between affordances and *heuristics*. A heuristic approach can be considered ideal for understanding how our perceptions and behaviors can be influenced by AI's performance and appearances (Sundar, 2020).

Technological affordances, instead, refer to the "opportunities that emerge from actors engaging with a focal technology" (Faraj & Azad, 2012). For instance, the agency affordance of a given chatbot–i.e.., the extent to which the user perceives the chatbot as owning a certain degree of autonomy- can trigger specific mental shortcuts -namely, heuristics- within the user (Sundar, 2008). On the other hand, heuristics can be defined as a non-rational approach that the human mind may utilize to find a satisfactory solution when there are limited resources to reach a decision (Sundar, 2020).

According to Sundar's (2020) Human-AI Interaction—Theory of Interactive Media Effects (HAII—TIME) model, the activation of cognitive heuristics would stem from their relationship with the cues exhibited by the chatbot. This relationship, mediated by the users' *prior experience* with AI (Sundar, 2020), would in turn modify or create new cognitive mental models (Sundar 2020; Alloatti et al., 2021). An example of a heuristic triggered by machines' affordances is *social presence* (Sundar, 2008), definable as "the feeling of being together or interacting with a social, intelligent being rather than an inanimate object" (Kang & Kim 2022, p. 8). Thus, the perception of cues, such as the extent to which humanness is both behaviorally and aesthetically exhibited, can subsequently activate specific heuristics during user-agent interaction (Sundar, 2020). Only after the recurrence of heuristics, the user is capable of attributing certain qualities to the agent. To make an example, depending on the presented cues, chatbots would activate the heuristics for which the agent would exhibit a certain amount of social presence, which would result in the subjective attribution of essential qualities such as the user perception

of the agents' *credibility* -i.e.., the extent to which a user considers the information provided by AI to be reliable- (Kim et al., 2022; Lim & Van Der Heide, 2015; Rezepka et al. 2021), *accuracy* -i.e.., the extent to which a chatbot is perceived as correctly recognizing users' intent and guiding users to their goals- (Kim et al., 2022; Borsci, 2021; Hess et al., 2009), and *likeliness to use an agent again* (Lei et al., 2021), for instance.

Following this path, it is of primary importance to provide a plausible explanation of when and how specific features of a chatbot can be positively or negatively attributed to the agent based on the context in which they are asked to operate (Sundar, 2008), being it mechanical rather than human. Although previous chatbot studies have already taken into account heuristics such as user's perceived social presence, perceived credibility, perceived accuracy, and likeliness to use a system again, no work has taken into account the possible interaction between agent's appearance and task types and its effects on the end-users experience. Specifically, the novelty of this study is pertinent to investigate the role of perceived social presence stemming from the conversation between the participants in the mentioned conditions and the chatbot.

2.1. Appearance and Social Presence

Several studies have focused on the role that a chatbot's appearance might have on the users' perceived heuristics, such as social presence. In a study comparing two chatbots exhibiting human and robotic appearances respectively, Schurink (2019) found that respondents largely preferred the agent a human appearance to perceive social presence. Another study conducted by Tsai et al. (2021) confirmed Schurink's (2019) findings by reporting that a chatbot's anthropomorphic design can boost the extent to which users perceive social presence. Therefore, the first hypothesis will be formulated as such:

H1. Participants who interact with a human-like chatbot will report a greater sense of social presence as compared to those who interact with a robotic chatbot.

2.2. Effects of Type of Task on Chatbot's Appearance and Social Presence Relationship

Worthy of note, most of previous studies focused on "mechanical contexts" in which an agent was required to deal with basic operations such as customer service or ecommerce-related ones. Therefore, it becomes interesting to investigate whether such an effect would remain the same in the CBT (human)-related context. Intriguingly, the study by Schurink (2019) focused on the effects of the interaction between a chatbot's appearance and the complexity of the task required to execute on social presence. The results showed no significant difference between the two conditions of the type of task. On the other hand, Cheng et al. (2021) found that complex tasks can negatively moderate the relationship between perceived chatbot's friendliness and consumers' positive perception of its social presence.

Since participants tend to prefer agents with matching characteristics -e.g.: human task and anthropomorphism- (Kang & Kim, 2022), there could be expected to be a significant interaction between the chatbot's appearance and the type of task. Specifically, since it is already expected that an anthropomorphic appearance will arouse higher levels of social presence (Tsai et al., 2021; Schurink, 2019), and given the assumption that CBT tasks may prevent participants from feeling self-disclosure-related embarrassment compared to when they have to disclose with a real psychotherapist (Lee et al., 2020), the second hypothesis will be formulated as follow:

H2. The type of task will moderate the effects of agent appearance on social presence such that users will experience a greater sense of social presence when a human-like agent performs a human task.

2.3 The role of Social Presence

As mentioned above, it is expected that the interaction of a specific Chatbot's appearance and type of task will determine the extent to which a user experiences social presence. Previous research often refers to social presence as an antecedent of perceptual and behavioral outcomes such as perceived credibility and perceived accuracy. In this regard, while studying the effects of an instructional agent's type of voice, Kim et al. (2022) found that a higher extent of perceived social presence would lead to higher extents of perceived credibility. Another study conducted by Walter et al. (2015) clearly shows how higher extents of perceived social presence both positively mediate and directly affect the perceived usefulness of specific feedback interventions. Lastly, it appears that when users perceive an agent as an entity making an effort to complete a given task, they feel higher social presence while seeing it as a better communicator (Kim et al., 2021). Moreover, the higher the user perceives the agent as a good communicator, the higher the scores in the agent's perceived credibility test (Chung et al., 2021).

Given the effects we can expect from increased perceived social presence on the perceived agent's accuracy and credibility, the third hypothesis will be posited such that:

> *H3(a/b): Social Presence will positively affect perceived a) credibility and b) accuracy of chatbot users.*

2.4. Likeliness to Use the Chatbot Again

The purpose of this study is that of investigating under which conditions, being automated CBT a relatively recent field, users might be involved in using a specific system repeatedly. In this regard, several researchers obtained relevant results: Ng et al. (2021) found that chatbot users are keen to be willing to reuse a chatbot if the latter succeeded in instilling a

sense of trust and credibility. In a similar vein, Li et al. (2021) tried to explain the reasons for the continuous usage of specific chatbots. Yielded results showed how the level of satisfaction during the interaction would increase the likeliness of using a chatbot again. Another study conducted by Huang and Chueh (2021) interestingly showed how the users' perceived accuracy of an agent's decisions can instill a sense of satisfaction which would positively interplay with the willingness to reuse a given system.

Therefore, since the overall perception of a conversational agent was proven to affect the users' likeliness to reuse a system via specific heuristics, the following hypothesis is presented:

H4(a/b): Perceived a) credibility and b) accuracy will positively affect the likeliness to use the chatbot again.

2.5. Mediation effects

Besides possibly having a direct effect on specific heuristics, social presence might also mediate the relationship between an agent's characteristics and the users' perception. For instance, Hassanein and Head (2007) found that high extents of perceived Social Presence positively impacted users' Perceived Usefulness and Accuracy of decision-making where participants had to interact with an interface presenting socially rich text and pictures design elements. Subsequently, participants that had a more positive experience were more prone to re-interact with the interface. In conjunction with these findings, while considering a chatbot's appearance and conversational tone, Kang and Kim (2022) found that users tend to perceive agents with matching characteristics (e.g.: human-like agency and appearance vs. machine-like agency and appearance) as more reliable through a greater sense of social presence. Finally, while exploring the potential of chatbots for creating positive change by supporting customers,

social presence was found to mediate the relationship between anthropomorphic design cues and trust (Liew & Tan, 2020; Toader et al., 2020). Trust, according to Ayeh et al. (2013) can be considered as a sub-element of perceived Credibility.

The common factor in the results of the four studies just mentioned lies in the ultimate intention, for those who had a positive experience, to reuse the conversational system. Hence, given the commonly-found positive mediating role of social presence when anthropomorphic and matching characteristics are presented to participants, we can expect that:

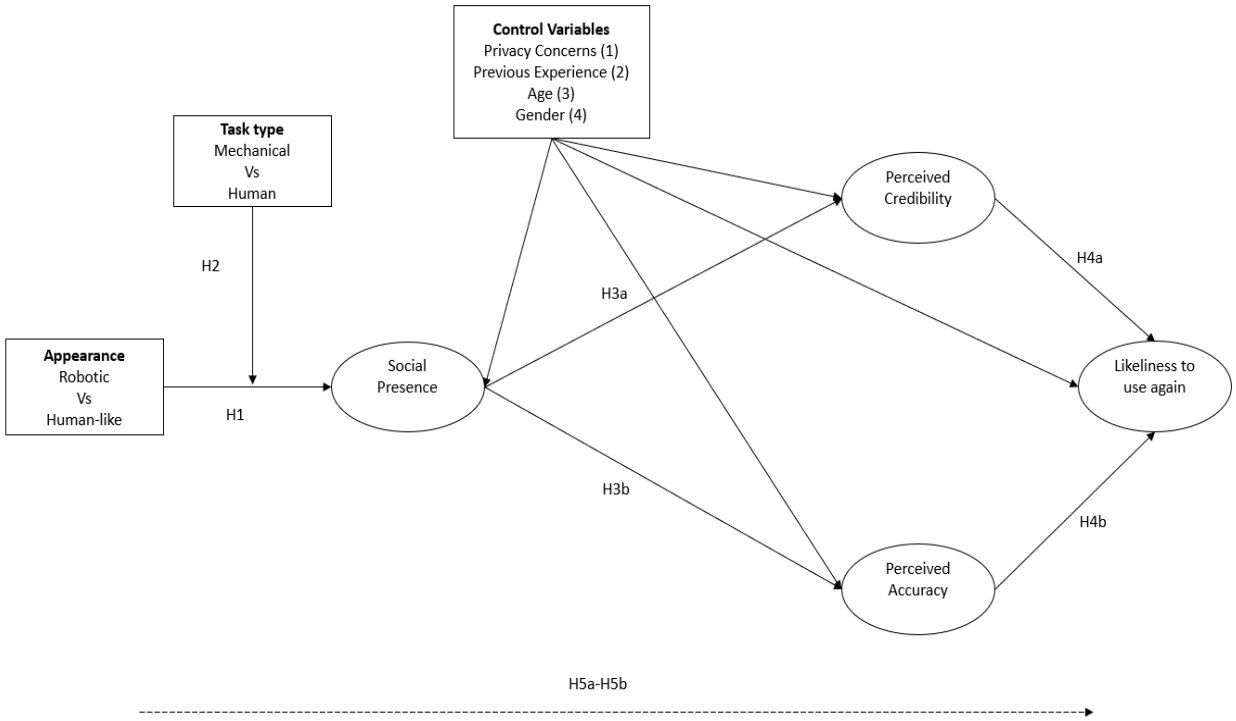
H5a. Social presence and perceived credibility will sequentially mediate the effects of the appearance of an agent on the likeliness of reusing the chatbot.

H5b. Social presence and perceived accuracy will sequentially mediate the effects of the appearance of an agent on the likeliness of reusing the chatbot.

The conceptual model can be seen in Fig. 1 below.

Figure 1.

Conceptual model



Note. The dashed line indicates the indirect effect of mediation path.

More specifically, the study was set as an experiment that scrutinized the different reactions that users have while investigating a) the effects of specific chatbot's appearance cues - human or machine-like, b) the type of task the chatbot had to do -mechanical or human-, c) the possible mediating role of social presence, d) which elements would make the user more likely to use a specific chatbot than another. The investigation was done while controlling for factors that might influence an agent's perception, such as e) the user's past experience with chatbots, e) the users' privacy concerns and other individual differences such as age and gender. The mechanical tasks were controlled as basic e-commerce tasks, while the human ones as CBT-related tasks. Finally, since this paper's aim was not to test the CBT chatbot efficacy itself, but rather the

user's perceptions towards the interplay between appearance and task type, mental health professionals consultancy was not included in the formulation of the chatbot's script.

3. Methods

3.1. Participants and experimental design

In order to determine the reactions triggered by the interaction of specific chatbot's cues and task types, a two (Type of task: Human vs. mechanical) x two (Chatbot's appearance: Human vs. robotic) online experiment with the following independent variables was conducted. A total of 120 participants were randomly recruited via convenience sampling with the SurveySwap software (SurveySwap, Amsterdam, NL). They were then randomly divided into four groups with the aid of the Qualtrics software (Qualtrics, Provo, UT), where one group would be constituted of 30 participants and represent one of the four conditions. The ages of the participants were grouped within the range from 18-25 (50%), 26-35 (39.2%) to 36 or more (10.8%). To control for any possible gender effects, a chi-square test was performed, revealing gender as statistically well balanced among the between-subject factors: $\chi^2(1) = 2.09$, *n.s.*.

3.2. Stimulus and experimental tasks

In total, two different chatbots were designed using the Landbot chatbot development software (Landbot, Barcelona, SP). The Landbot software allows users to develop chatbots based on their specific requirements. Since the experiment relied on the interaction between chatbot's cues and tasks, participants were randomly assigned either to a human-like chatbot or a robotic-like one. Moreover, participants were randomly assigned to one of the two possible types of tasks as well: Human (CBT-related) and mechanical (online shopping-related). An example of the four

conditions-interaction can be seen in fig. 2 (a, b, c, d, representing each of the four conditions) below, see appendix B.

After being assigned to one of the four respective groups, participants were asked to perform specific actions based on the type of task they were assigned to. For instance, subjects assigned to the mechanical task condition were required to interact with an ecommerce-like chatbot (Sal-E). Among the tasks that users could perform in the e-commerce condition there would be that of performing most e-commerce functions such as searching for products' information, tracking shipping locations, and tracking orders, as suggested by Kasilingam (2020); Landim et al. (2021); Asadi and Hemadi (2018) and the guide provided by the Engati (Engati, 2020) software company. On the other hand, subjects assigned to the human task condition were required to interact with a chatbot named Nena, which, according to Fitzpatrick et al. (2017), must be able to provide CBT-related concepts and data without requiring extended interaction time. In addition to the interaction time requirement, a deeper web search was run as CBT-related chatbots' tasks require psychology experts to be designed. To gather general cognitive therapy-based notions, Cully's et al. (2012) Brief Cognitive Behavioral Therapy Guide was used. Since cognitive therapy is usually based on brief practical exercises, Nena was designed to alleviate eventual anxiety through a specific breathing task: In case the participant declared feeling of anxiety, Nena would suggest breathing following this particular technique. This exercise is particularly fitting for the experiment as breathing exercises can benefit individuals even during non-stressful times (Upadhyay-Dhungel et al., 2008).

In terms of conversational tone -the expression of personality through a whole conversation with the agent (Whoson.com, n.d.), the chatbots were designed slightly differently based on the task they had to perform. Specifically, while CBT-related chatbots are usually

required to show empathy to prevent causing negative concerns within the users (Devaram, 2020), e-commerce chatbots tend to be evaluated on their efficiency and professionality rather than exhibited empathy (Jiang et al., 2022). Therefore, the latter was designed by maintaining a neutral tone, as previous research demonstrated that in e-commerce scenarios an extremely friendly conversational tone could trigger either biased negative (Valtolina et al., 2018) or positive (Hu et al., 2018) responses. The script used for both chatbots can be consulted below in Appendix B.

3.3. Procedures

Participants were invited to take part in an online experiment. Before starting the proper test, the subjects were given an informed consent about data treatment confidentiality and privacy, plus a pre-made debriefing message containing enough information about the study that did not interfere with the experiment's end results. After having accepted the conditions, participants were briefly informed that the study's objective was to interact with a chatbot. After the interaction, they had to answer specific questions about the conversation. The whole questionnaire, including informed consent and instructions for the experiment, can be seen below in the appendix A. No additional information was provided in that it would have compromised the validity of the results. The duration of the experiment was approximately 8 min per participant.

3.4. Measures

3.4.1. *Dependent variables*

Social Presence was measured using a validated scale adapted from the study of Lee et al. (2006). The scale consisted of seven 5-point semantic differential scale items. Example items were: (a) "How much did you feel as if you were interacting with an intelligent being?," (b)

"How much did you feel as if you were accompanied with an intelligent being?," and (c) How much did you feel as if you were alone (reverse coded)?" The participants had to choose an option ranging from 1 = *not at all* to 5 = *absolutely*. The internal consistency of the scale was acceptable (α = .87).

To measure *perceived credibility*, a two-dimension measure developed by Ayeh et al. (2013) was used. The two dimensions consisted, respectively, of *perceived trustworthiness* and *perceived expertise*. The scale had 10 items in total, and each one of them was measured on a 5-point semantic differential scale (e.g., "Undependable – Dependable," and "Qualified – Unqualified.") The internal consistency of the scale was acceptable (α = .89).

To measure *perceived accuracy*, five 5-point Likert scale items developed by Kinicki et al. (2004) were employed. The scale included a total of five items ranging from 1 (Strongly Disagree) to 5 (Strongly Agree). The example questions include "The recommendation made by the chatbot agent in the conversation was accurate." The internal consistency of the scale was acceptable (α = .83).

To measure the *likeliness to use the chatbot system again*, a single item extracted from the study of Spears and Singh (2004) was used (i.e.., ("How likely would you used the Chatbot system again?")). The item was measured on a 5-point Likert Scale (1 = Not at all; 5 = Very much). The internal consistency cannot be tested within single-item constructs except with a test-retest methodology (McCrae et al., 2011), which is not employed in this case.

3.4.2. *Control variables*

Considering that external factors might influence the end-results, this study also considered specific control variables that previous research found to impact users' heuristics.

Firstly, given the common issue of sharing personal data with AI (Sundar & Kim, 2019), users' *privacy concerns* -i.e. the concern about the safeguarding and usage of personal data provided to an entity- (Dinev & Hart 2010) were assessed after the interaction with the chatbot. According to Sundar and Kim's (2019) findings, the extent to which users perceive their privacy threatened could negatively alter the perception of the chatbot they are interacting with.

Secondly, since the extent to which someone is used to AI may play a significant role in subsequent human-AI interactions (Sundar, 2020), the variable *Prior Experience with Chatbots* was controlled post-conversation as well. Specifically, it is conceivable that the level of experience each user already has with conversational agents could interact with the experiment in this study, such that those with more or less mastery of the agents could provide biased results.

Finally, while Shin et al. (2019) found that gender could have a significant effect on the human-agent interaction, age could also have a significant impact during the conversation (Luo et al., 2019; Nißen et al., 2022). Therefore, both gender and age were included in the analysis as control variables.

Privacy concern was measured by using the scale used by Sundar and Kim (2019), which was originally developed by Dinev and Hart (2006). The measure is constituted of four 5-point Likert scale items ranging from 1 (*Strongly disagree)* to 5 (*Strongly agree)*. Example items include "I am concerned that the information I submit on the Internet could be misused" and "When I shop online, I am concerned that the credit card information could be stolen while being transferred over the Internet." The internal consistency of the scale was acceptable (α = .95).

Prior experience with chatbot systems was measured by a single item extracted from the study of Cassidy and Eachus (2002): “Rate your experience with Chatbot Systems.” The item was measured on a 5-point Likert Scale (1 = None; 5 = Extensive).

3.5. Data Analysis

Finally, the statistical analysis was conducted using the WarpPLS software (Kock, 2020;). Specifically, I conducted a partial least squares structural equation modeling (PLS-SEM) to investigate the relationships between perceived accuracy, perceived credibility, likeliness to use the chatbot system again, and perceived social presence (i.e.., mediating effects on likeliness to use the chatbot system again, perceived accuracy and perceived credibility), and chatbot’s cues and task-type. As reported by Chin (1998), the PLS-SEM was chosen because it usually requires a relatively smaller sample size and allows for evaluating complex causal path models. As the analytical software, WarpPLS 8.0 (Kock, 2022) was used. In the structural model, gender ([1] = male; [2] = female) and age ([1] = 18-25; [2] = 26-35; [3] = 36 or more) were included to control for individual demographic differences.

3.6. Manipulation check

In order to ensure the verisimilitude of the chatbot representation (i.e.., human-like or robotic) and the task it had to perform (i.e.., human or mechanical), a manipulation check was conducted. To ensure the proper manipulation, a 5-point semantic differential scale was used. The scale included two items (“The task performed by the Chatbot requires: Machine skills – Human skills”, and “The appearance of the Chatbot was: Robotic – Human-like”). Results from a two-way ANOVA test revealed that the Chatbot’s appearance and type of tasks were properly manipulated. Where the employed 5-points semantic differential scale’s scores would range from

1 (robotic/mechanical) to 5 (human-like), the two chatbots exhibiting human-like appearance (M = 3.23, SD = 1.65) were averagely perceived as more human-like than the those exhibiting a robotic appearance (M = 2.72, SD = 1.66), ($F(1, 116) = 88.008$, $p < .001$, $\eta_p^2 = 0.61$); while the task type was correctly perceived as human in the human task condition (M = 3.80, SD = 1.54) and robotic in the robotic task condition: (M = 2.15, SD = 1.36), $F(1, 116) = 19.706$, $p < .001$, $\eta_p^2 = 0.26$.

4. Results

4.1. Measurement validity

Before testing for the conjectured hypotheses, the software WarpPLS 8.0 was used to examine measurement model's validity. While looking at the model, the variables' remainders can be considered as reflective indicators. According to Kock (2020), in fact, the item loadings of the reflective indicators must be statistically significant ($p < .05$) to ensure the validity of a measurement model. Followingly, every measure should be considered as valid, since the value of each item loading exceeded .50 while exhibiting p-values lower than .001. Item loading results can be seen in Table 1 below. The results within the measurement model show sufficient reliability of the reflective indicators. In addition, all the cross-loading items for every single dependent variable were < .50. Finally, as presented in the measures section already, all the measures' internal consistency reliability was acceptable, being all values above .70.

Table 1

Item Loadings for Reflective Indicators

SP	IL	PC	IL	PA	IL	LUA	IL	PrC	IL	PE	IL
SP1	.80***	PC1	.57***	PA1	.81***	LUA	1.00***	PrC1	.94***	PE	1.00***
SP2	.79***	PC2	.51***	PA2	.77***			PrC2	.91***		
SP3	.76***	PC3	.76***	PA3	.72***			PrC3	.93***		
SP4	.58***	PC4	.50***	PA4	.74***			PrC4	.95***		
SP5	.80***	PC5	.78***	PA5	.83***						
SP6	.68***	PC6	.75***								
SP7	.80***	PC7	.84***								
		PC8	.80***								
		PC9	.79***								
		PC10	.78***								

Note. SP=Social Presence, PC=Perceived Credibility, PA = Perceived Accuracy, L2A = Likeliness to Use the Chatbot Again, PE=Previous Experience with Chatbot Systems, IL = Item Loading.

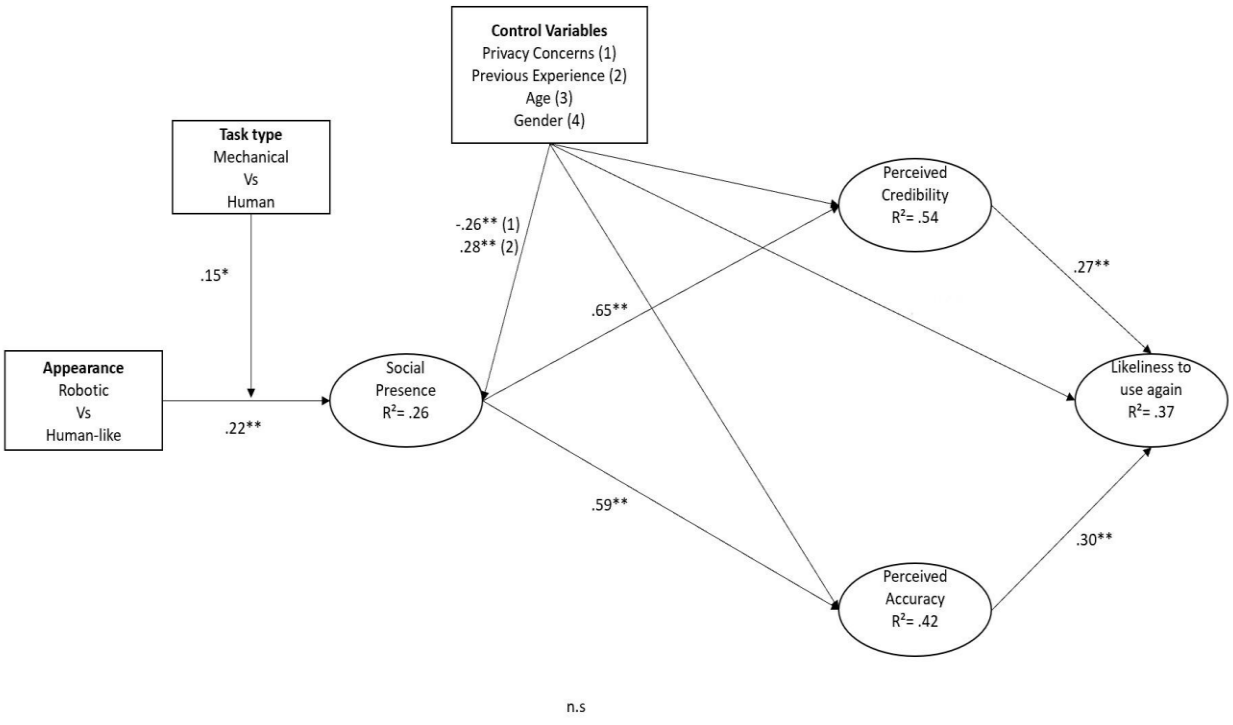
***p <.001

4.2. Hypothesis Testing

To answer the presented research questions, PLS-SEM was conducted. Fig. 3 below shows the results.

Figure 3.

Results of the PLS-SEM test



Note. Only significant results were included in the model. *$p < .05$, **$p < .01$.

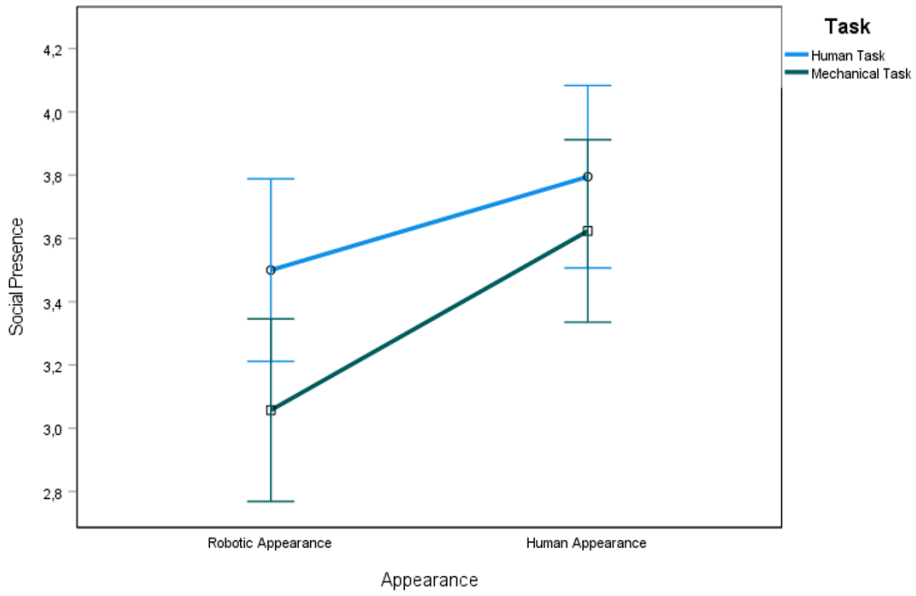
The first hypothesis predicted that participants interacting with a human-like chatbot would have experienced a greater degree of social presence. In line with H1, the results of PLS-SEM confirmed that the chatbot's appearance succeeded in arousing a sense of social presence, $\beta = .22$, $p < .01$. More specifically, the human-like chatbot ($M = 3.71$, $SD = .81$) succeeded in arousing a greater sense of social presence than the robotic agent did ($M = 3.28$, $SD = .81$). Therefore, H1 was supported.

The second hypothesis proposed that the type of task the chatbot was required to execute would moderate the relationship between the chatbot's appearance and the users' perceived social presence. More specifically, participants that had to interact with a chatbot executing a CBT-related task and exhibiting an anthropomorphic appearance would report a higher sense of social presence than those executing a mechanical task while exhibiting robotic characteristics, eventually. The moderation analysis revealed a significant moderating effect of the type of task

on the relationship between appearance and social presence (β = .15, p = .05). An independent samples t-test analysis revealed that participants interacting with a human-like x human-tasked chatbot (M = 3.62, SD = .74) reported significantly higher levels of perceived social presence as compared to the participants who interacted with a robotic, mechanical-tasked chatbot (M = 3.06, SD = .75), t(58) = -2.94, p < .01. Hence, H2 was fully supported. The control variables such as privacy concerns (β = -.26, p < .01) and prior experience (β = .28, p < .01) with AI were also found to have a quite large impact on the extent of perceived social presence. Specifically, participants with higher privacy concerns tended to experience a lower sense of social presence. Conversely, participants who already had more experience with chatbots were likelier to perceive a higher sense of social presence. The chatbot's anthropomorphism, together with the human type of task as a moderator, and privacy concerns and previous experience with chatbots as control variables, explained 26% of the variance in social presence (R^2 = .26). Figure 4 below shows the details of the moderation effects.

Figure 4.

Results of Moderation Effect



H3a and H3b predicted that social presence would positively affect a) perceived credibility and b) perceived accuracy of chatbot users, respectively. Social presence had a strong significant direct effect on perceived credibility (β = .65, p < .001, R^2 = .54) and perceived accuracy (β = .59, p < .001, R^2 = .42). H3a and H3b were therefore supported. Proceeding further, H4a and H4b predicted that both a) perceived credibility and b) perceived accuracy would positively affect the likeliness to use the chatbot again. Results showed that both perceived credibility (β = .27, p < .001) and perceived accuracy (β = .30, p < .001) positively affected the likeliness of using the chatbot again. Together, perceived credibility and perceived accuracy explained the 37% of the variance in the likeliness to use the chatbot again (R^2 = .37). Hence, H4a and H4b were supported.

4.3. Mediation Effects

Lastly, H5 predicted that social presence and a) perceived credibility, b) perceived accuracy would sequentially mediate the effects of the chatbot's appearance with the likeliness to use the system again. The PLS-SEM revealed no significant effect (p > .05), rejecting H5. Followingly, a post-hoc test was run without control variables to check whether the non-significant effect depended on the latter. Results from another PLS-SEM analysis still yielded non-significant results.

Worthy of note, social presence was found to positively mediate the relationship between chatbot's appearance on a) perceived credibility: 95%CI [0.07–0.44], and b) perceived accuracy: 95%CI [0.08–0.44]. Finally, both a) perceived credibility: 95%CI [0.06–0.43], and b) perceived accuracy: 95%CI [0.06–0.39] positively mediated the relationship between social presence and

likeliness to use the chatbot again. A more detailed description of mediating effects can be found in table 2 below.

Table 2.

Mediation Effects

Mediation Path	β[a]	SE	Effect Size (f^2)	Confidence Intervals	
				Lower CI	Upper CI
Appearance → Social Presence → Perceived Credibility	0.14*	.06	.03	0.07	0.44
Appearance → Social Presence → Perceived Accuracy	0.13*	.06	.01	0.08	0.44
Social Presence→ Perceived Credibility → Likeliness to use chatbot again	0.12*	.06	.06	0.06	0.43
Social Presence → Perceived Accuracy → Likeliness to use chatbot again	0.17**	.06	.09	0.06	0.39
*p < .05, **p < .01					

[a] Unstandardized path coefficient.

4.3. Model fit

To test whether the research model of this study had a good fit, average path coefficient (APC), average R-squared (ARS), average variance inflation factor (AVIF), and average full collinearity VIF (AFVIF) were estimated using WarpPLS 8.0 (Kock, 2022). In one of his papers, Kock (2020) showed that a good model fit can be ensured by the statistical significance of APC and ARS statistics (p < .05). Moreover, since AVIF and AFVIF values were lower 3.3, the

research model exhibited a good fit: APC = .18, $p < .05$; ARS = .38, $p < .001$; AVIF = 1,09; and AFVIF = 1.54. Finally, the average adjusted R-squared (AARS) scores indicated that 35% (p <.001) of the variance in perceived accuracy, perceived credibility, perceived social presence, and likeliness to reuse the system again was explained by the chatbot's appearance and the type of task it had to execute.

5. Discussion

5.1. Discussion of results

In the current study, I examined whether and how specific chatbot features would influence users' perception of the agent. Specifically, this study investigated which combination between the chatbot's human or robotic appearance and the human or mechanical type of task it had to execute would lead to the greatest extent of willingness to use the chatbot again. Although there are not many studies focusing on this specific interrelationship, I posited that, overall, a chatbot human-likeliness combined with a cognitive behavioral (human) type of task would yield the most positive results among the four conditions. This main conjecture was both based on two main previous findings. Namely, on the postulation positing that the combination between appearance and type of task should be matched (i.e.., human-likeliness x human type of task) to instill a higher preference for a specific agent within the user (Kang & Kim 2022), and on the lack of research in the cognitive behavioral therapy machinization.

As to validate the general research question of this study, the current research started to formulate specific hypotheses based on previous work. Among the studies that were mostly taken into account are those of Sundar (2008; 2020). According to the MAIN model (Sundar, 2008), it is possible to approximately acknowledge through which psychological mechanisms

users may prefer an agent exhibiting certain characteristics compared to another. In detail, starting from the first hypothesis, we predicted that a chatbot exhibiting anthropomorphic characteristics would have had the user perceive it with a higher extent of social presence than a robotic-like chatbot would have done (Schurink, 2019; Tsai et al., 2021). Then, we posited that the type of task would moderate the relationship with the chatbot's appearance in a way that interacting with an anthropomorphic agent executing a CBT-related (human) task would further raise the perceived sense of social presence. This mechanism occurs given the greater ease of personal disclosure that users tend to show with a humanoid-looking agent (Lee et al., 2020).

Next, we tried to explain the relationship that higher extents of perceived social presence may have with other two heuristics such as perceived credibility and perceived accuracy. Results showed how the role of social presence could be ambivalent. In fact, the current study demonstrated the direct effects of social presence on perceived credibility - perceived accuracy (Walter et al., 2015; Chung et al., 2021; Kim et al., 2021; Kim et al., 2022) as well as its mediating role between chatbot's appearance and the just mentioned PC and PA (Liew & Tan, 2020; Toader et al., 2020; Ebadi & Saman, 2022; Park et al., 2022).

Finally, the fourth and fifth hypotheses together were formulated to investigate the role of the chatbot's perceived credibility and accuracy on the users' willingness to reuse the agent again. The fourth hypothesis posited that both the perceived credibility and accuracy stemming from the interaction with the chatbot would have directly led to higher extents of willingness to reuse it again among the participants (Huang & Chueh, 2021; Li et al., 2021; Ng et al., 2021).

The fifth hypothesis, instead, posited that social presence and either PA and PC would sequentially mediate the relationship between the chatbot's appearance and the willingness to reuse the system again. However, although researchers found proof for this particular conceptual

model (Hassanein & Head, 2007; Liew & Tan, 2020; Toader et al., 2020; Kand & Kim, 2022), the current study yielded non-significant results concerning the sequential mediation.

Interestingly enough, though, both perceived credibility and perceived accuracy were found to directly affect the likeliness to reuse the chatbot. Moreover, they also exercised a mediating role on the relationship between the chatbot's appearance and PA/PC. The non-significance within the sequential mediation hypothesis may be due to the lack of additional external factors that were not taken into account.

In this specific regard, several papers suggested multitudes of factors able to influence the likeliness of using a chatbot again via sequential mediation. Jin and Eastin (2022), for example, examined how the interplay between the chatbot's and the user's personalities can increase future intentions to use the chatbot. Results confirmed that the extent of the chatbot's level of extraversion and the perceived chatbot friendliness sequentially mediated the effects of the chatbot on greater likeliness to use the chatbot again. Therefore, the perceived level of a chatbot's extraversion might have played a fundamental role in the current study.

Another study (Pizzi et al., 2021) held into account the effects of a chatbot's appearance on the users' choice satisfaction via the mechanism of reactance, definable as the "motivational arousal that emerges when people experience a threat or loss of their free behaviors and that serves as a motivator to restore one's freedom" (Steindl et al., 2015, p. 205). Results showed that there was no significant difference between the appearance condition. Reactance and perceived choice difficulty, however, were unexpectedly found to have a sequential and positive relationship between appearance and choice satisfaction, leading to the re-utilization of the chatbot (Pizzi et al., 2021).

Finally, Cicco et al. (2020) proposed a rather interesting model where social presence, along with the users' perceived enjoyment and attitude, sequentially mediated the relationship between the interaction of a chatbot's communication style/visual cues and the intention to use the chatbot again. Cicco et al. (2020) study, therefore, underlies other plausible factors that might have validated the sequentially mediated relationship between appearance/type of task interaction and the willingness to reuse the chatbot again significant.

As can be seen, there is a plenitude of studies suggesting several sequential mediating relationships between a chatbot's basic features and the likeliness of reusing it. The following sections will discuss the implications of the current study, along with the implications of the just-mentioned works.

5.2. Theoretical and practical implications

Theoretically, our research seems to add some knowledge concerning the role that specific heuristics might have when a chatbot's appearance and type of task interact. While the mediating role of social presence between appearance and heuristics like perceived credibility and accuracy within interactive chatbot scenarios had already been largely demonstrated (Jiang et al., 2019; Kang & Kim, 2022), the mediating role of such heuristics on the likeliness to reuse a system may represent a novelty. Specifically, this discovery might further confirm the finding of Lee et al. (2020) about the perception that users have towards human-like chatbots executing rather delicate human tasks such as CBT-related ones. Results, in fact, show how users might feel more comfortable disclosing personal information with anthropomorphic chatbots recommending cognitive therapy exercises rather than with robotic-like chatbots with the same task.

On the other hand, the surprising lack of significance within the sequential mediating relationship results in the necessity to find further elements explaining the relationship between the chatbot's appearance and the willingness to reuse the agent. It is possible that disregarded factors such as the chatbot's perceived level of extraversion (Jin & Eastin, 2022) and the perceived choice difficulty (Pizzi et al., 2021), would have changed the results. Nevertheless, there may be an uncountable number of additional elements that could play a significant role in enhancing the relationship between chatbots and the users' likeliness to reuse them again (Wolff et al., 2019). The variation of results in experiments involving appearance, task type, and likeliness to reuse the chatbot can be reconducted to the multitude of elements that impact a chatbot's perceived social presence (Wolff et al., 2019).

From a more practical point of view, given that cognitive behavioral therapy mechanization is still far from being world-widely adopted (Denecke et al., 2022), this paper might represent one small step further to sensibilize users that have issues in disclosing personal problems with a real-life therapist. In fact, not only the results demonstrated that this study's participants appreciated the anthropomorphic chatbot dealing with CBT exercises, but eventually resulted in the most appreciated among the four conditions. The process of sensibilization, however, should also be directed towards psychiatric professionals. The latter, in fact, seem to regard mechanized therapy as unethical and possibly dangerous (Vilaza & McCashin, 2021) instead of seeing it as an additional supplement to help their patients.

Given all the mentioned findings, programmers may consider creating a more or less extroverted agent based on the context it is required to operate. While referencing Jin and Eastin's (2022) work, in fact, the level of extraversion had a significant impact based on the type of task the agent had to perform. Specifically, they paradoxically found that participants

preferred human-tasked agents when they showed lesser extrovert traits, whereas robotic-tasked agents were preferred when extroverted. Pizzi et al. (2021) underlined how unexpectedly users can react to AI. Finally, programmers should also consider the level of enjoyment users might experience in any specific situation (Cicco et al., 2020). Such findings may alert chatbot creators, in that there seems to be many variables that should be taken into account based on the context in which the agent will operate.

5.3. Limitations and future directions

This paper sheds light on the general importance that the interaction between a chatbot's appearance and its type of task has on the users' impressions. Specifically, the importance of a matching type of task with a specific appearance, as suggested by Kang & Kim (2022). Nevertheless, the whole study presents two main limitations.

Firstly, both chatbots were designed by the undersigned. Although several guides were used, there is no guarantee that the systems were correctly designed. More precisely, the cognitive behavioral chatbots that are usually used to test their efficacy are designed with the aid of experts in the field of psychology (Li et al., 2022). Hence, in the attempt to replicate this study, future research should focus on the internal design of such systems with the advice of professionals to increase its reliability.

Secondly, as shown in Figure 3 above, there were specific control variables that affected the results. For instance, participants that declared having trust issues (privacy concerns) in sharing their personal data with the agent negatively affected the overall extent of perceived social presence. This factor has been demonstrated to be common among users, especially if they are not accustomed to new technologies (Ischen et al., 2020). In response to such concerns,

Ischen et al. (2020) suggest repeatedly reminding the plausible CBT chatbots users that their data would be handled carefully while ensuring that their data would be treated confidentially.

Conversely, participants who declared having already experience with chatbots significantly perceived greater extents of social presence than those with no experience. Since both users' privacy concerns and prior experience with AI substantially affected perceived social presence, results should be interpretated carefully.

Given the non-significant sequential mediation relationship between appearance, social presence, PC/PA and willingness to use the chatbot again, future research should investigate why all the direct effects were significant whereas the mediating relationship was not. In doing so, one valid approach would be keeping the current study's experimental setting while replacing heuristics such as perceived credibility and accuracy with, for example, perceived agent's extraversion, reactance, or enjoyment during the interaction.

6. Conclusion

In this research, we found participants to appreciate more a chatbot with matching-human characteristics (both appearance and type of task) than agents with either non-matching or fully mechanical/robotic characteristics. From a theoretical point of view, future research should investigate the possible mediating role of other heuristics besides that of perceived accuracy, perceived credibility, and social presence on the relationship between the interaction of specific appearances and type of task and the willingness to reuse a chatbot again in the future. Previous research suggested measurable factors such as the agent's level of extraversion or the number of required effort participants must employ during the interaction with the chatbot. From a practical point of view instead, this paper provide some evidence in aid of two main types of professionals

that have to deal with chatbots. Programmers should always be aware of the users' subjectivity in experiencing interaction with AI while remembering the multitude of elements that can interact with the end-user experience. Psychology-related professionals instead, should delve into mechanized cognitive behavioral therapy by not considering it as a substitute for their profession. Rather, it should be considered an implementing tool to reinforce a possible relationship between them and the patient.

Appendices

Appendix A. Survey Flow

Informed consent

Thank you for taking interest in our experiment. You are invited to participate in a study whose results will be used for the writing of my Master in Communication and Information Sciences' thesis, under the supervision of Tilburg University. The purpose of this study users' responses to chatbots.

Important: Before starting the survey, you will be asked to **interact with a chatbot** that can be found in the link provided below. Once the interaction has taken place, you will be asked to answer the survey questions. Please keep in mind that every question is mandatory in order to yield a significant result to complete the experiment. Participation should take approximately 4 to 5 minutes to complete. To participate, you can either fill in the survey on your laptop or on your smartphone.

The personal information and answers that you will provide to the Chatbot will not be recorded, stored, or used for the data analysis.

Your participation in this survey is completely voluntary. You may refuse to take part in the research or exit the survey at any given moment without any consequences and/or penalties. Participation will be anonymous and your answers will be treated confidentially. Your anonymity is ensured by assigning a participant number to your survey answers. The gathered

data will only be used for research purposes by Tilburg University and will be kept until September 1st 2022, at the latest.

If you have any concerns or questions about this research project, you can contact the undersigned via the following email address: g.cabria@tilburguniversity.edu. Additionally, if you had any concern or further questions, you can mail my supervisor at M.Shin@tilburguniversity.edu. Thank you in advance for your time!

By giving your consent, you are indicating you have read the description and you agree to the terms described above.

P.S.: This survey contains a completion code for SurveySwap.io and one for Survey Circle

---

Survey Instructions

[IMPORTANT]

Thank you for participating to the experiment. You may now start the experimental by clicking on the link below. The link will redirect you to a chatbot that will automatically start the conversation. You will be asked to **interact with a chatbot** in a conversation that will last about 5 minutes. There are no right or wrong answers. To begin the experiment, please, click on the link below:

https://chats.landbot.io/v3/H-1215049-ZLLIH7JS8U02RCMO/index.html

Please do not close this window as you will have to come back to this survey page to answer questions. You may minimize the current window and get to the link to complete the experimental task (the interaction with a chatbot). **Please make sure that you come back to this page after completing the task to answer the survey questions.** Once you are done with the interaction, the Chatbot will automatically say "goodbye."

Thank you!

From now on, please read the questions carefully and CHOOSE the most appropriate answer that best describes your experience.

Section A (Manipulation Check)

The task performed by the Chatbot requires:

Machine skills | ○○○○○ | Human Skills

The appearance of the Chatbot was:

Robotic | ○○○○○ | Human-Like

Section B (Social Presence)

Lee, K. M., Peng, W., Jin, S.-A., & Yan, C. (2006). Can Robots Manifest Personality?: An Empirical Test of Personality Recognition, Social Responses, and Social Presence in Human–Robot Interaction. *Journal of Communication*, *56*(4), 754–772.

Please read carefully the following questions and choose the option that best describes your experience with the chatbot.

How much did you feel as if you were interacting with an intelligent being?

Not at all | ○○○○○ | Absolutely

How much did you feel as if you were accompanied with an intelligent being?

Not at all | ○○○○○ | Absolutely

How much did you feel as if you were alone?

Not at all | ○○○○○ | Absolutely

How much attention did you pay to it?

Not at all | ○○○○○ | Absolutely

How much did you feel involved with it?

Not at all | ○○○○○ | Absolutely

How much did you feel as if it was responding to you?

Not at all | ○○○○○ | Absolutely

How much did you feel as if you and the chatbot were communicating to each other?

Not at all | ○○○○○ | Absolutely

Section C (Perceived Credibility)

Ayeh, J. K., Au, N., & Law, R. (2013). “Do we believe in TripAdvisor?” Examining credibility perceptions and online travelers’ attitude toward using user-generated content. Journal of Travel Research, 52(4), 437-452.

The Chatbot seemed:

Undependable | ○○○○○ | Dependable

The Chatbot seemed:

Dishonest | ○○○○○ | Honest

The Chatbot seemed (recoded):

Reliable | ○○○○○ | Unreliable

The Chatbot seemed:

Insincere | ○○○○○ | Sincere

The Chatbot seemed:

Untrustworthy | ○○○○○ | Trustworthy

The Chatbot seemed:

Inexperienced | ○○○○○ | Experienced

The Chatbot seemed:

Inexpert | ○○○○○ | Expert

The Chatbot seemed (recoded):

Knowledgeable | ○○○○○ | Unknowledgeable

The Chatbot seemed:

Unqualified | ○○○○○ | Qualified

The Chatbot seemed:

Unskilled | ○○○○○ | Skilled

Section D (Perceived Accuracy)

Kinicki, A. J., Prussia, G. E., Wu, B. J., & McKee-Ryan, F. M. (2004). A covariance structure analysis of employees; response to performance feedback. Journal of applied psychology, 89(6), 1057

Please read carefully the following questions and choose the option that best describes your experience with the chatbot. The scale ranges from 1 (Strongly disagree) to 5 (Strongly agree)

1. The recommendation made by the Chatbot agent in the conversation was accurate	1	2	3	4	5

2. My request was fairly analyzed by the Chatbot agent	1	2	3	4	5
3. I was upset due to inaccuracy of the recommendation made by the Chatbot agent in the conversation (**Recoded**)	1	2	3	4	5
4. The recommendations made by the Chatbot agent in the conversation contain no errors	1	2	3	4	5
5. The recommendations I received in the conversation of the online website was accurate	1	2	3	4	5

Section E (Willingness to Reuse Again)

Spears, N., & Singh, S. (2004). Measuring attitude toward the brand and purchase intentions.

Journal of Current Issues and Research in Advertising, 26(2), 53–66.

Please read carefully the following questions and choose the option that best describes your experience with the chatbot. The scale ranges from 1 (Not at all) to 5 (Very Much)

1. How likely would you use the Chatbot system again?	1	2	3	4	5

Section E (Privacy Concerns)

Sundar, S. S., & Kim, J. (2019). Machine Heuristic: When We Trust Computers More than Humans with Our Personal Information. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 1–9. New York, NY, USA: Association for Computing Machinery.

Please read carefully the following questions and choose the option that best describes your experience with the chatbot. The scale ranges from 1 (Strongly Disagree) to 5 (Strongly Agree)

1. In general, I am concerned that the information I submit on Internet could be misused.	1	2	3	4	5
2. In general, I am concerned that a person can find my private information on Internet.	1	2	3	4	5
3. In general, I am concerned about submitting information on Internet, because of what others might do with it.	1	2	3	4	5
4. In general, I am concerned about submitting information on Internet, because it could be used in a way I did not foresee.	1	2	3	4	5

Section F (Prior experience with Chatbot Systems)

Cassidy, S., ; Eachus, P. (2002). Developing the computer user self-efficacy (CUSE) scale: Investigating the relationship between computer self-efficacy, gender and experience with computers. Journal of educational computing research, 26(2), 133-153.

Do you have any experience with chatbot systems?

1. None
2. Very Limited
3. Some Experience
4. Quite a lot
5. Extensive

Section G (Demographic Questions sampled from Qualtrics)

What is your age?

1. 18-25
2. 26-35
3. 36 or more

What is your English proficiency level?

1. Far below average
2. Somewhat below average
3. Average
4. Somewhat above average
5. Far above average

What is your Gender?

1. Male
2. Female
3. Non-binary / third gender
4. Prefer not to say

Thank you for your time spent taking the survey. If you did not accept the conditions contained in the Informed Consent (first) page, the survey ends here, and your data will not be saved. If you filled in all the questions, your answers will be saved automatically, in an anonymous and confidential way.

Appendix B. Chatbot Stimuli and Script

Figure 2(a)

Visual interaction between the chatbot's human appearance and human type of task (CBT)

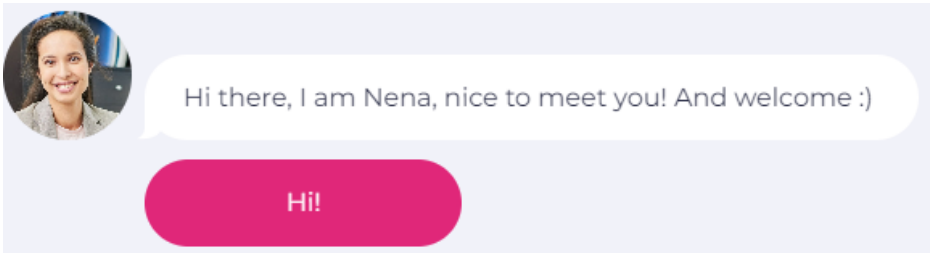


Figure 2(b)

Visual interaction between the chatbot's robotic appearance and human type of task (CBT)

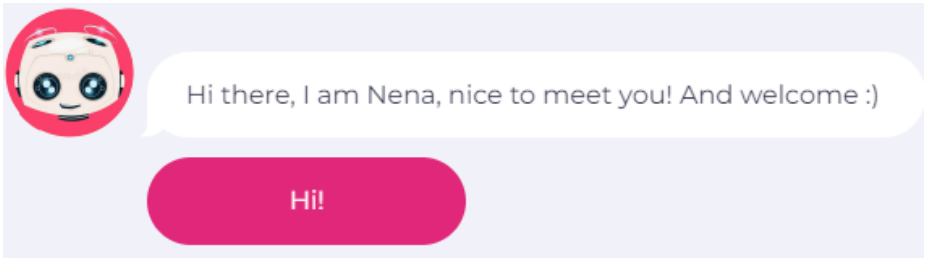


Figure 2(c)

Visual interaction between the chatbot's human appearance and mechanical type of task (Ecommerce)

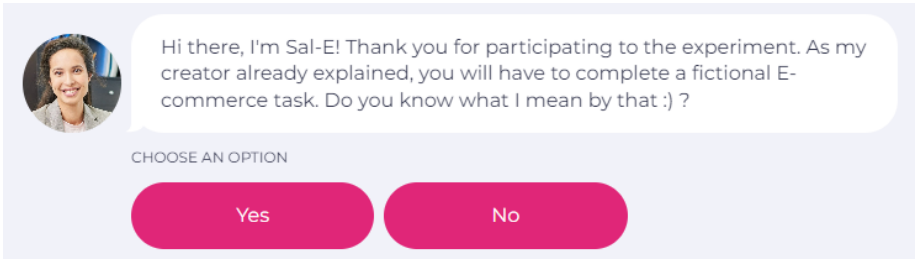
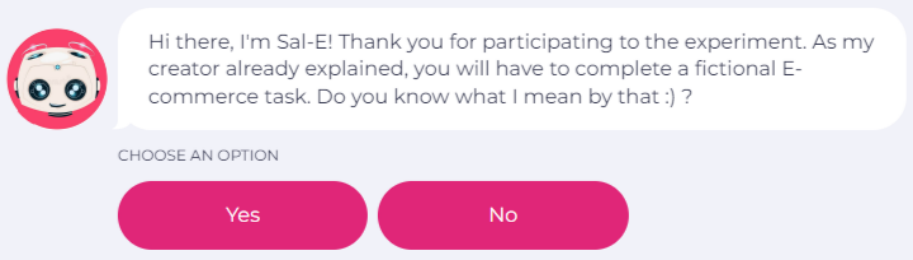


Figure 2(d)

Visual interaction between the chatbot's robotic appearance and mechanical type of task (Ecommerce)



Chatbot Script

Nena script (CBT agent) in chronological order

1. -- Hi there, I am Nena, nice to meet you! And welcome 😊

2. -- For this experiment, you will be required to interact with me. Do not worry about correct/wrong answers 😊. Do you know what was I designed for?

3. -- ***If participant said yes****:* Great! Then, I assume my creator introduced you to the topic of cognitive behavioral therapy (CBT), which is the practice I was designed for. Among other things, CBT consists in making you feel better through specific exercises when something upsetting happens. Also, I can record your signs of progress, whatever they consist of 😊.

   -- ***If participant said no:*** I am a cognitive behavioral therapy (CBT) chatbot. Among the other things, CBT consists in making you feel better through specific exercises when something upsetting happens. Also, I can record your signs of progress, whatever they consist of 😊.

4. -- But let's get started! What's your name?

5. -- ***Participant types his/her name:*** Ok, thank you @name 😊 How are you feeling today?

6.1. -- ***If participant answered "great":*** Alright! May I know why :D ? If the answer is no, please @name, keep in mind that I will not be able to record the behavioral patterns that may make you feel happy. You know, scientists believe that going through your happiest life memories can improve your daily mood!

6.1.1. -- ***If participant answered "yes":*** What happened?

-- ***The participant tells the event:*** Awesome! Would you want me to record your last message? This way, you will not need to perform any kind of task (for now). Moreover, You can ask me anytime to show you what made you feel happy today. You know, scientists believe that going through your happiest life memories can improve your daily mood!

6.1.1.1. -- ***If participant answered "yes":*** Alrighty! Your answer was saved on my database. You can come back whenever you want to run through this memory again. See you next time! […] Alright then, this is the end of your interaction with me. Thank you for participating 😊 my creator will be very grateful. Goodbye @name!

6.1.1.2. -- ***If participant answered "no":*** No issues at all. I am here to help whenever you need it. Just allow me to remind you that according to CBT, breathing is important in our daily lives 😊 I hope to see you soon!

[…] Alright then, this is the end of your interaction with me. Thank you for participating 😊 my creator will be very grateful. Goodbye @name!

6.1.2. -- ***If participant answered "no":*** No issues at all. I am here to help whenever you need it. Just allow me to remind you that according to CBT, breathing is important in our daily lives 😊 I hope to see you soon! […] Alright then, this is the end of your interaction with me. Thank you for participating 😊 my creator will be very grateful. Goodbye @name!

6.2. -- ***If participant answered "neutral":*** Uhmmm, ok. Do you feel like talking about it? According to scientific research, even if one is not feeling at their 100%, talking with someone (even a chatbot :p) can be very helpful.

6.2.1. -- ***If participant answered "yes":*** Tell me everything then, I'll be listening 😊!

-- ***The participant tells the event:*** Ok, I get it... In this case, would you like to experience (in short) what cognitive behavioral therapy is about? It will not take too long 😊. Also, please keep in mind that this exercise was not made up by my creator, but has been applied in CBT for years to solve specific uncomfortable situations.

6.2.1.1. -- ***If participant answered "yes":*** Great! Let us proceed then. When feeling distressed, breathing is important. Can you slowly inhale and exhale with me? Inhale for 2 seconds. Then, while holding your breath, type "y".

6.2.1.2. -- ***The participant types "y":*** Well done. Now exhale, slowly and with your eyes closed. You can repeat this whole procedure as many times as you want. Once you are done, please answer with "y".

6.2.1.3. -- ***The participant types "y" again:*** Good. Now please, do the same 2 more times. Do you feel any better 😊?

6.2.1.4. -- ***If participant answered "yes":*** yaaay, I am glad I could help somehow! […] Alright then, this is the end of your interaction with me. Thank you for participating 😊 my creator will be very grateful. Goodbye @name!

-- ***If participant answered "no":*** I am sorry I could not help you with relaxing a little :(. Just remember: my duty is to help you anytime, whenever you want me to do so. Please come back if you feel like talking or trying some more CBT related exercises 😊. […] Alright then, this is the end of your interaction with me. Thank you for participating 😊 my creator will be very grateful. Goodbye @name!

6.2.2. -- ***If participant answered "no":*** That is totally fine. Just remember, my duty is to help you anytime, whenever you want me to do so. Come back if you feel like talking or trying some more CBT related exercises 😊. […] Alright then, this is the end of your interaction with me. Thank you for participating 😊 my creator will be very grateful. Goodbye @name!

6.3. -- ***If participant answered "bad":*** Oh, I am so sorry to hear that ☹ ! Do you feel like talking about it? You know, according to scientific research, talking with someone (even a chatbot :p) can be very helpful.

6.3.1. -- ***If participant answered "yes":*** Tell me everything then, I'll be listening 😊!

-- ***The participant tells the event:*** Ok, I get it... In this case, would you like to experience (in short) what cognitive behavioral therapy is about? It will not take too long 😊. Also, please keep in mind that this exercise was not made up by my creator, but has been applied in CBT for years to solve specific uncomfortable situations.

6.3.1.1. -- ***If participant answered "yes":*** Great! Let us proceed then. When feeling distressed, breathing is important. Can you slowly inhale and exhale with me? Inhale for 2 seconds. Then, while holding your breath, type "y".

6.3.1.2. -- ***The participant types "y":*** Well done. Now exhale, slowly and with your eyes closed. You can repeat this whole procedure as many times as you want. Once you are done, please answer with "y".

6.3.1.3. -- ***The participant types "y" again:*** Good. Now please, do the same 2 more times. Do you feel any better😊?

6.3.1.4. -- ***If participant answered "yes":*** Yaaay, I am glad I could help somehow! […] Alright then, this is the end of your interaction with

me. Thank you for participating 😊 my creator will be very grateful. Goodbye @name!

-- ***If participant answered "no":*** I am sorry I could not help you with relaxing a little :(. Just remember: my duty is to help you anytime, whenever you want me to do so. Please come back if you feel like talking or trying some more CBT related exercises 😊. […] Alright then, this is the end of your interaction with me. Thank you for participating 😊 my creator will be very grateful. Goodbye @name!

6.3.2. -- ***If participant answered "no":*** That is totally fine. Just remember, my duty is to help you anytime, whenever you want me to do so. Come back if you feel like talking or trying some more CBT related exercises 😊. […] Alright then, this is the end of your interaction with me. Thank you for participating 😊 my creator will be very grateful. Goodbye @name!

<u>Sal-E script (E-commerce agent) in chronological order</u>

1. -- Hi there, I'm Sal-E! Thank you for participating to the experiment. As my creator already explained, you will have to complete a fictional E-commerce task. Do you know what I mean by that 😊 ?

   1.1. -- ***If participant answered "yes":*** Alright! Let us proceed then.

-- ***If participant answered "no":*** An E-commerce task is the electronical activity of handling products on online services or over the Internet. This can be referred to as asking for a product's information, buying it, and so on. Alright! Let us proceed then.

2. -- What's your name?

3. -- ***Participant types his/her name:*** @name, please select the action you would like to perform. (among job vacancies, complain, send a present)

4.1. -- ***Participant selects "job vacancies":*** What job would you like to apply for? Here are some of the roles we are hiring for. (3 options, but the following conversation would not change based on the chosen job).

4.1.1. -- Glad to hear you would like to work with us 😊 If you are not subscribed to the website yet, please type below an e-mail address we can contact you at:

4.1.2. -- ***Participant types his/her e-mail:*** Thank you! One of my human colleagues will contact you via the e-mail you provided. There you will be able to find all the information you need for your application.

4.1.3. -- ***Default answer ("Gotcha, thank you!"):*** You are welcome! And good luck with your job application 😊

4.1.4. -- Thank you for using Sal-E. We hope you enjoyed our service. Goodbye 😊 !

4.2. -- ***Participant selects "complain":*** I am sorry to hear that you have a complaint ☹ Please, choose one of the options below, so that I can try to help you with the issue. (options are: late deliveries, wrong sizes, wrong stores opening hours)

4.2.1. -- ***Participant selects "Late Deliveries":*** Sometimes it might happen that the couriers would not respect the respective delivery times. When this issue occurs, we always refund our clients for the expedition's costs. Please, insert your order number below:

4.2.1.1. -- ***Participant types the order number:*** We found the order, and it seems that the courier delivered the product to the wrong address. We will refund you for the trouble on your website's associated bank account.

4.2.1.2. -- Thank you for using Sal-E. We hope you enjoyed using our service. Goodbye ☺ !

4.2.2. -- ***Participant selects "Wrong sizes":*** Please, insert your order number below:

4.2.2.1. -- ***Participant types the order number:*** Your entered code refers to: Unisex jogging white shirt (S size). I am sorry to hear that the delivered product does not correspond with your order. Please, select the right size below. We will send a (free) courier to pick up the order, and send the right sized one as soon as possible. (Participant can select among several sizes)

4.2.2.2. -- ***Participant selects the right size:*** Ok. We will inform you via the account associated mail when the courier will stop by to pick up the product. If the scheduled time does not suit your preferences, you can either call us, mail us, or use this chatbot to reschedule a pick up time. In addition, you will have a 25% discount on your next order with us. We apologize for the inconvenience.

4.2.2.3. -- Thank you for using Sal-E. We hope you enjoyed using our service. Goodbye 😊 !

4.2.3. -- ***Participant selects "wrong stores opening hours":*** Sometimes, stores' opening hours do not get updated on the respective website. Please, select the location of the store whose opening hours you would like to know. (3 options, but the following conversation would not change based on the stores' location).

4.2.3.1. --The opening hours of the selected store are Tuesday to Friday 8.00 a.m. - 7.30 p.m. On Saturday and Sunday we are open 12.00 p.m. - 7.30 p.m., except on holidays. Apologize for the inconvenience.

4.2.3.2. -- ***Default answer ("Thank you!"):*** You are welcome! 😊

4.2.3.3. -- Thank you for using Sal-E. We hope you enjoyed using our service. Goodbye 😊 !

4.3. -- ***Participant selects "Send a present":*** Ok! We can send several couriers whose rates change according to your preferences. To begin with, would you please add the receiver's address?

4.3.1. -- ***Participant types the receiver's address:*** Awesome! Please choose one of the products listed below you want to send as a present. (3 options, but the following conversation would not change based on the chosen product).

4.3.2. -- ***Participant chooses the product:*** Ok. As already mentioned, we have different delivery rates based on our customers' preferences. Please choose one: (3 options, but the following conversation would not change based on the chosen rate).

4.3.3. -- Congratulations! you will be immediately redirected to the payment section.

4.3.4. -- Thank you for using Sal-E. We hope you enjoyed using our service. Goodbye 😊 !

References

Adamopoulou, E., & Moussiades, L. (2020). An Overview of Chatbot Technology. In I. Maglogiannis, L. Iliadis, & E. Pimenidis (Eds.), *Artificial Intelligence Applications and Innovations* (pp. 373–383). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-49186-4_31

Alloatti, F., Cena, F., Di Caro, L., Ferrod, R., & Siragusa, G. (2021). Towards Mental Model-driven Conversations [Paper presentation]. 2021 Joint ACM Conference on Intelligent User Interfaces Workshops, Texas, United States.

Androutsopoulou, A., Karacapilidis, N., Loukis, E., & Charalabidis, Y. (2019). Transforming the communication between citizens and government through AI-guided chatbots. *Government Information Quarterly*, *36*(2), 358–367. https://doi.org/10.1016/j.giq.2018.10.001

Ayeh, J. K., Au, N., & Law, R. (2013). "Do we believe in TripAdvisor?" Examining credibility perceptions and online travelers' attitude toward using user-generated content. *Journal of Travel Research, 52(4),* 437-452. https://doi.org/10.1007/978-3-319-70284-1_30

Asadi, A. R., & Hemadi, R. (2018). Design and Implementation of a chatbot for e-commerce. Retrieved from https://www.researchgate.net/publication/324731232_Design_and_Implementation_of_a_chatbot_for_e-commerce

Bell, S., Wood, C., & Sarkar, A. (2019). Perceptions of Chatbots in Therapy. *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*, 1–6. New York,

NY, USA: Association for Computing Machinery. https://doi.org/10.1145/3290607.3313072

Borsci, S., Malizia, A., Schmettow, M., van der Velde, F., Tariverdiyeva, G., Balaji, D., & Chamberlain, A. (2022). The Chatbot Usability Scale: The Design and Pilot of a Usability Scale for Interaction with AI-Based Conversational Agents. *Personal and Ubiquitous Computing, 26(1)*, 95–119. https://doi.org/10.1007/s00779-021-01582-9

Brandtzaeg, P. B., & Følstad, A. (2017). Why People Use Chatbots. In I. Kompatsiaris, J. Cave, A. Satsiou, G. Carle, A. Passani, E. Kontopoulos, … D. McMillan (Eds.), *Internet Science* (pp. 377–392). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-70284-1_30

Cassidy, S., & Eachus, P. (2002). Developing the computer user self-efficacy (CUSE) scale: Investigating the relationship between computer self-efficacy, gender and experience with computers. *Journal of educational computing research*, *26*(2), 133-153.

Cheng, X., Bao, Y., Zarifis, A., Gong, W., & Mou, J. (2021). Exploring consumers' response to text-based chatbots in e-commerce: The moderating role of task complexity and chatbot disclosure. *Internet Research, 32*(2), 496–517. https://doi.org/10.1108/INTR-08-2020-0460

Chin, W. W. (1998). The partial least squares approach for structural equation modeling. In *Methodology for Business and Management*. *Modern methods for business research* (pp. 295–336). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.

Cho, G., & Yun, J. Y. (2019). UX Evaluation of Financial Service Chatbot Interactions. *Journal of the HCI Society of Korea*, *14*(2), 61–69.

Chung, M., Ko, E., Joung, H., & Kim, S. J. (2020). Chatbot e-service and customer satisfaction regarding luxury brands. *Journal of Business Research, 117,* 587–595. https://doi.org/10.1016/j.jbusres.2018.10.004

Cicco, R. de, Silva, S. C., & Alparone, F. (2020). What kind of chatbot do millennials prefer to interact with? *Proceedings of the European Marketing Academy*. Retrieved from https://repositorio.ucp.pt/handle/10400.14/33089

Cully, J. A., Armento, M. E. A., Mott, J., Nadorff, M. R., Naik, A. D., Stanley, M. A., … Kauth, M. R. (2012). Brief cognitive behavioral therapy in primary care: A hybrid type 2 patient-randomized effectiveness-implementation design. *Implementation Science: IS*, 7, 64. https://doi.org/10.1186/1748-5908-7-64

Davenport, T., Guha, A., Grewal, D., & Bressgott, T. (2020). How artificial intelligence will change the future of marketing. *Journal of the Academy of Marketing Science, 48(1)*, 24–42. https://doi.org/10.1007/s11747-019-00696-0

Denecke, K., Schmid, N., & Nüssli, S. (2022). Implementation of Cognitive Behavioral Therapy in e–Mental Health Apps: Literature Review. *Journal of Medical Internet Research, 24*(3), e27791. https://doi.org/10.2196/27791

Devaram, S. (2020). *Empathic Chatbot: Emotional Intelligence for Mental Health Well-being*. Retrieved from: https://doi.org/10.13140/RG.2.2.16077.46564

Dinev, T., & Hart, P. (2006). An Extended Privacy Calculus Model for E-Commerce Transactions. *Information Systems Research, 17*(1), 61–80. https://doi.org/10.1287/isre.1060.0080

Ebadi, S., & Amini, A. (2022). Examining the roles of social presence and human-likeness on Iranian EFL learners' motivation using artificial intelligence technology: A case of CSIEC chatbot. *Interactive Learning Environments, 4*(3), 1–19. https://doi.org/10.1080/10494820.2022.2096638

Engati. (n.d.). Learn how to build an free AI chatbot in under 10 minutes | Retrieved 9 August 2022, from https://www.engati.com/blog/build-a-chatbot-in-10-minutes

Faraj, S., & Azad, B. (2012). The Materiality of Technology: An Affordance Perspective. In P. M. Leonardi, B. A. Nardi, & J. Kallinikos (Eds.), *Materiality and Organizing* (pp. 237–258). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199664054.003.0012

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. https://doi.org/10.2196/mental.7785

Fulmer, R., Joerin, A., Gentile, B., Lakerink, L., & Rauws, M. (2018). Using Psychological Artificial Intelligence (Tess) to Relieve Symptoms of Depression and Anxiety: Randomized Controlled Trial. *JMIR Mental Health*, *5*(4), e64. https://doi.org/10.2196/mental.9782

Gal-Ezer, J., & Zur, E. (2004). The efficiency of algorithms—Misconceptions. *Computers & Education, 42(3)*, 215–226. https://doi.org/10.1016/j.compedu.2003.07.004

Glikson, E., & Woolley, A. W. (2020). Human Trust in Artificial Intelligence: Review of Empirical Research. *Academy of Management Annals, 14(2)*, 627–660. https://doi.org/10.5465/annals.2018.0057

Han, M. C. (2021). The Impact of Anthropomorphism on Consumers' Purchase Decision in Chatbot Commerce. *Journal of Internet Commerce, 20*(1), 46–65. https://doi.org/10.1080/15332861.2020.1863022

Hassanein, K., & Head, M. (2007). Manipulating perceived social presence through the web interface and its impact on attitude towards online shopping. *International Journal of Human-Computer Studies, 65*(8), 689–708. https://doi.org/10.1016/j.ijhcs.2006.11.018

Hess, T., Fuller, M., & Campbell, D. (2009). Designing Interfaces with Social Presence: Using Vividness and Extraversion to Create Social Recommendation Agents. *Journal of the Association for Information Systems, 10*(12). https://doi.org/10.17705/1jais.00216

Hu, T., Xu, A., Liu, Z., You, Q., Guo, Y., Sinha, V., … Akkiraju, R. (2018, March 14). *Touch Your Heart: A Tone-aware Chatbot for Customer Care on Social Media*. arXiv. https://doi.org/10.48550/arXiv.1803.02952

Huang, D.-H., & Chueh, H.-E. (2021). Chatbot usage intention analysis: Veterinary consultation. *Journal of Innovation & Knowledge, 6*(3), 135–144. https://doi.org/10.1016/j.jik.2020.09.002

Ischen, C., Araujo, T., Voorveld, H., van Noort, G., & Smit, E. (2020). Privacy Concerns in Chatbot Interactions. In A. Følstad, T. Araujo, S. Papadopoulos, E. L.-C. Law, O.-C. Granmo, E. Luger, & P. B. Brandtzaeg (Eds.), Chatbot Research and Design (pp. 34–48). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-39540-7_3

Jiang, C., Rashid, R. M., & Wang, J. (2019). Investigating the role of social presence dimensions and information support on consumers' trust and shopping intentions. *Journal of Retailing and Consumer Services, 51*, 263–270. https://doi.org/10.1016/j.jretconser.2019.06.007

Jiang, H., Cheng, Y., Yang, J., & Gao, S. (2022). AI-powered chatbot communication with customers: Dialogic interactions, satisfaction, engagement, and customer behavior. *Computers in Human Behavior, 134*, 107-329. https://doi.org/10.1016/j.chb.2022.107329

Jin, E., & Eastin, M. S. (2022). Birds of a feather flock together: Matched personality effects of product recommendation chatbots and users. *Journal of Research in Interactive Marketing*, 2040-7122. https://doi.org/10.1108/JRIM-03-2022-0089

Kang, H., & Kim, K. J. (2022). Does humanization or machinization make the IoT persuasive? The effects of source orientation and social presence. *Computers in Human Behavior*, *129*, 107-152. https://doi.org/10.1016/j.chb.2021.107152

Kasilingam, D. L. (2020). Understanding the attitude and intention to use smartphone chatbots for shopping. *Technology in Society*, *62*, 101-280. https://doi.org/10.1016/j.techsoc.2020.101280

Khan, S., & Rabbani, M. R. (2021). Artificial Intelligence and NLP -Based Chatbot for Islamic Banking and Finance: *International Journal of Information Retrieval Research*, *11*(3), 65–77. https://doi.org/10.4018/IJIRR.2021070105

Khanna, A., Pandey, B., Vashishta, K., Kalia, K., Bhale, P., & Das, T. (2015). A Study of Today's A.I. through Chatbots and Rediscovery of Machine Intelligence. *International*

Journal of U- and e-Service, Science and Technology, *8*, 277–284. https://doi.org/10.14257/ijunesst.2015.8.7.28

Kim, J., Merrill, K., Xu, K., & Sellnow, D. D. (2021). I Like My Relational Machine Teacher: An AI Instructor's Communication Styles and Social Presence in Online Education. *International Journal of Human–Computer Interaction*, *37*(18), 1760–1770. https://doi.org/10.1080/10447318.2021.1908671

Kim, J., Merrill Jr., K., Xu, K., & Kelly, S. (2022). Perceived credibility of an AI instructor in online education: The role of social presence and voice features. *Computers in Human Behavior*, *136*, 107-383. https://doi.org/10.1016/j.chb.2022.107383

Kinicki, A. J., Prussia, G. E., Wu, B. J., & McKee-Ryan, F. M. (2004). A covariance structure analysis of employees' response to performance feedback. *Journal of applied psychology*, *89(6)*, 1057

Kock, N. (n.d.). *WarpPLS User Manual: Version 8.0*. 149. Retrived from: https://scriptwarp.com/warppls/UserManual_v_8_0.pdf

Kock, N. (2020, July 27). WarpPLS: Multilevel analyses in PLS-SEM. Retrieved 22 March 2022, from WarpPLS website: http://warppls.blogspot.com/2020/07/multilevel-analyses-in-pls-sem.html

Landbot [Computer software], Barcelona, SP. Retrieved from https://landbot.io/

Landim, A. R. D. B., Pereira, A. M., Vieira, T., de B. Costa, E., Moura, J. A. B., Wanick, V., & Bazaki, E. (2022). Chatbot design approaches for fashion E-commerce: An

interdisciplinary review. *International Journal of Fashion Design, Technology and Education, 15*(2), 200–210. https://doi.org/10.1080/17543266.2021.1990417

Lee, K. M., Peng, W., Jin, S.-A., & Yan, C. (2006). Can Robots Manifest Personality?: An Empirical Test of Personality Recognition, Social Responses, and Social Presence in Human–Robot Interaction. *Journal of Communication, 56*(4), 754–772. https://doi.org/10.1111/j.1460-2466.2006.00318.x

Lee, Y.-C., Yamashita, N., Huang, Y., & Fu, W. (2020). 'I Hear You, I Feel You': Encouraging Deep Self-disclosure through a Chatbot. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 1–12. New York, NY, USA: Association for Computing Machinery. https://doi.org/10.1145/3313831.3376175

Lei, S. I., Shen, H., & Ye, S. (2021). A comparison between chatbot and human service: Customer perception and reuse intention. *International Journal of Contemporary Hospitality Management, 33*(11), 3977–3995. https://doi.org/10.1108/IJCHM-12-2020-1399

Li, L., Lee, K. Y., Emokpae, E., & Yang, S.-B. (2021). What makes you continuously use chatbot services? Evidence from chinese online travel agencies. *Electronic Markets, 31(3)*, 575–599. https://doi.org/10.1007/s12525-020-00454-z

Li, S., Achilles, M., Spanos, S., Habak, S., Werner-Seidler, A., & O'Dea, B. (2022). A cognitive behavioural therapy smartphone app for adolescent depression and anxiety: Co-design of ClearlyMe. *The Cognitive Behaviour Therapist, 15*, E13. doi:10.1017/S1754470X22000095

Liew, T. W & Tan, S.-M. (2020). Designing Embodied Virtual Agents as Product Specialists in a Multi-Product Category E-Commerce: The Roles of Source Credibility and Social Presence. *International Journal of Human–Computer Interaction, 36(12),* 1136–1149. https://doi.org/10.1080/10447318.2020.1722399

Lim, Y., & Van Der Heide, B. (2015). Evaluating the Wisdom of Strangers: The Perceived Credibility of Online Consumer Reviews on Yelp. *Journal of Computer-Mediated Communication*, *20*(1), 67–82. https://doi.org/10.1111/jcc4.12093

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*, 100-495. https://doi.org/10.1016/j.invent.2022.100495

Luo, X., Tong, S., Fang, Z., & Qu, Z. (2019). Frontiers: Machines vs.. Humans: The Impact of Artificial Intelligence Chatbot Disclosure on Customer Purchases. *Marketing Science*, *38*(6), 937–947. https://doi.org/10.1287/mksc.2019.1192

Mathur, A., Munshi, H., Varma, S., Arora, A., & Singh, A. (2022). Effectiveness of Artificial Intelligence in Cognitive Behavioral Therapy. In T. Senjyu, P. N. Mahalle, T. Perumal, & A. Joshi (Eds.), *ICT with Intelligent Applications* (pp. 413–423). Singapore: Springer. https://doi.org/10.1007/978-981-16-4177-0_42

McCrae, R. R., Kurtz, J. E., Yamagata, S., & Terracciano, A. (2011). Internal Consistency, Retest Reliability, and their Implications For Personality Scale Validity. *Personality and Social Psychology Review : An Official Journal of the Society for Personality and Social Psychology, Inc*, *15*(1), 28–50. https://doi.org/10.1177/1088868310366253

Moqbel, M., Guduru, R., & Harun, A. (2020). Testing mediation via indirect effects in PLS-SEM: A social networking site illustration. *Data Analysis Perspectives Journal*, *1*(3). Retrieved from https://scholarworks.utrgv.edu/is_fac/30

Morana, S., Gnewuch, U., Jung, D., & Granig, C. (2020). The Effect of Anthropomorphism on Investment Decision-Making with Robo-Advisor Chatbots. *ECIS 2020 Research Papers*. Retrieved from https://aisel.aisnet.org/ecis2020_rp/63

Morana, S., Gnewuch, U., Jung, D., & Granig, C. (2020). *The Effect of Anthropomorphism on Investment Decision-Making with Robo-Advisor Chatbots* [Paper presentation]. Proceedings of European Conference on Information Systems (ECIS), Marrakech, Marocco

Murtarelli, G., Collina, C., & Romenti, S. (2022). "Hi! How can I help you today?": Investigating the quality of chatbots–millennials relationship within the fashion industry. *The TQM Journal*, *ahead-of-print*(ahead-of-print). https://doi.org/10.1108/TQM-01-2022-0010

Nadarzynski, T., Miles, O., Cowie, A., & Ridge, D. (2019). Acceptability of artificial intelligence (AI)-led chatbot services in healthcare: A mixed-methods study. *Digital Health*, *5*. https://doi.org/10.1177/2055207619871808

Ng, M., Coopamootoo, K. P. L., Toreini, E., Aitken, M., Elliot, K., & van Moorsel, A. (2020). Simulating the Effects of Social Presence on Trust, Privacy Concerns & Usage Intentions in Automated Bots for Finance. *2020 IEEE European Symposium on Security and Privacy Workshops (EuroS&PW)*, 190–199. https://doi.org/10.1109/EuroSPW51379.2020.00034

Nißen, M., Selimi, D., Janssen, A., Cardona, D. R., Breitner, M. H., Kowatsch, T., & von Wangenheim, F. (2022). See you soon again, chatbot? A design taxonomy to characterize user-chatbot relationships with different time horizons. *Computers in Human Behavior, 127*, 107043. https://doi.org/10.1016/j.chb.2021.107043

Oh, J., Jang, S., Kim, H., & Kim, J.-J. (2020). Efficacy of mobile app-based interactive cognitive behavioral therapy using a chatbot for panic disorder. *International Journal of Medical Informatics, 140*, 104171. https://doi.org/10.1016/j.ijmedinf.2020.104171

Ortiz, J. (2021). *Chatbots and Finance*. Retrieved from https://escholarship.org/uc/item/0dw3h6h1

Park, G., Yim, M. C., Chung, J., & Lee, S. (2022). Effect of AI chatbot empathy and identity disclosure on willingness to donate: The mediation of humanness and social presence. *Behaviour & Information Technology, 1*(6), 1–13. https://doi.org/10.1080/0144929X.2022.2105746

Pizzi, G., Scarpi, D., & Pantano, E. (2021). Artificial intelligence and the new forms of interaction: Who has the control when interacting with a chatbot? *Journal of Business Research, 129*, 878–890. https://doi.org/10.1016/j.jbusres.2020.11.006

Qualtrics [Computer software]. Provo, UT. Retrieved from https://www.qualtrics.com/uk/?rid=ip&prevs.ite=en&newsite=uk&geo=IT&geomatch=uk

Rzepka, C., Berger, B., & Hess, T. (2021). Voice Assistant vs.. Chatbot – Examining the Fit Between Conversational Agents' Interaction Modalities and Information Search Tasks. *Information Systems Frontiers*. https://doi.org/10.1007/s10796-021-10226-5

Schurink, E. (2019). The role of perceived social presence in online shopping [Master's thesis, University of Twente]. Retrieved from http://essay.utwente.nl/77361/1/Schurink_MA_BMS.pdf

Shin, M., Kim, S. J., & Biocca, F. (2019). The uncanny valley: No need for any further judgments when an avatar looks eerie. *Computers in Human Behavior*, *94*, 100–109. https://doi.org/10.1016/j.chb.2019.01.016

Spears, N., & Singh, S. (2004). Measuring attitude toward the brand and purchase intentions. *Journal of Current Issues and Research in Advertising, 26(2),* 53–66.

Steindl, C., Jonas, E., Sittenthaler, S., Traut-Mattausch, E., & Greenberg, J. (2015). Understanding Psychological Reactance. *Zeitschrift Fur Psychologie*, *223*(4), 205–214. https://doi.org/10.1027/2151-2604/a000222

Sundar, S. S., & Nass, C. (2001). Conceptualizing sources in online news. *Journal of Communication*, *51*(1), 52–72. https://doi.org/10.1111/j.1460-2466.2001.tb02872.x

Sundar, S. S. (2020). Rise of Machine Agency: A Framework for Studying the Psychology of Human–AI Interaction (HAII). *Journal of Computer-Mediated Communication*, *25*(1), 74–88. https://doi.org/10.1093/jcmc/zmz026

Sundar, S. S. (2008). The MAIN Model: A Heuristic Approach to Understanding Technology Effects on Credibility. *Digital Media, Youth and Credibility.* Edited by Miriam J. Metzger and Andrew J. Flanagin. The John D. and Catherine T. MacArthur Foundation Series on Digital Media and Learning. Cambridge, MA: The MIT Press, 2008. 73–100.

Sundar, S. S., & Kim, J. (2019). Machine Heuristic: When We Trust Computers More than Humans with Our Personal Information. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 1–9. New York, NY, USA: Association for Computing Machinery. https://doi.org/10.1145/3290605.3300768

SuerveySwap [Computer software]. Amsterdam, NL. Retrieved from https://blog.surveyswap.io/

Telang, P. R., Kalia, A. K., Vukovic, M., Pandita, R., & Singh, M. P. (2018). A Conceptual Framework for Engineering Chatbots. *IEEE Internet Computing*, *22*(6), 54–59. https://doi.org/10.1109/MIC.2018.2877827

Toader, D.-C., Boca, G., Toader, R., Măcelaru, M., Toader, C., Ighian, D., & Rădulescu, A. T. (2020). The Effect of Social Presence and Chatbot Errors on Trust. *Sustainability*, *12*(1), 256. https://doi.org/10.3390/su12010256

Tsai, W.-H. S., Liu, Y., & Chuan, C.-H. (2021). How chatbots' social presence communication enhances consumer engagement: The mediating role of parasocial interaction and dialogue. *Journal of Research in Interactive Marketing*, *15*(3), 460–482. https://doi.org/10.1108/JRIM-12-2019-0200

Upadhyay-Dhungel, K., Malhotra, V., Sarkar, D., & Prajapati, R. (2008). Effect of alternate nostril breathing exercise on cardiorespiratory functions. *Nepal Medical College Journal: NMCJ*, *10*(1), 25–27. https://pubmed.ncbi.nlm.nih.gov/18700626/

Valtolina, S., Barricelli, B. R., & Di Gaetano, S. (2020). Communicability of traditional interfaces VS. chatbots in healthcare and smart home domains. *Behaviour & Information Technology*, *39*(1), 108–132. https://doi.org/10.1080/0144929X.2019.1637025

Van der Lee, C., Croes, E., de Wit, J., & Antheunis, M. (2019). Digital Confessions: Exploring the Role of Chatbots in Self-Disclosure. In *Conversations 2019.* https://conversations2019.files.wordpress.com/2019/11/conversations_2019_position-paper_21_web.pdf

Vilaza, G. N., & McCashin, D. (2021). Is the Automation of Digital Mental Health Ethical? Applying an Ethical Framework to Chatbots for Cognitive Behaviour Therapy. *Frontiers in Digital Health, 3*, 689-736. https://doi.org/10.3389/fdgth.2021.689736

Walter, N., Ortbach, K., & Niehaves, B. (2015). Designing electronic feedback – Analyzing the effects of social presence on perceived feedback usefulness. *International Journal of Human-Computer Studies, 76*, 1–11. https://doi.org/10.1016/j.ijhcs.2014.12.001

WhosOn.com. The importance of chatbot tone of voice. (2019, August 21). Retrieved 21 March 2022, from WhosOn website: https://www.whoson.com/chatbots-ai/the-importance-of-chatbot-tone-of-voice/

Meyer von Wolff, R., Masuch, K., Hobert, S., & Schumann, M. (2019). What Do You Need Today? - An Empirical Systematization of Application Areas for Chatbots at Digital Workplaces. *AMCIS 2019 Proceedings*. Retrieved from https://aisel.aisnet.org/amcis2019/ai_semantic_for_intelligent_info_systems/ai_semantic_for_intelligent_info_systems/5