

Consumer Perception of Chatbots and Purchase Intentions: Anthropomorphism and Conversational Relevance

¹Sooyun Iris Chung, ²Kwang-Hee Han

¹Masters Student, Department of Psychology, Yonsei Univ., Korea

²*Prof., Department of Psychology, Yonsei Univ., Korea
sooyuniris@yonsei.ac.kr, khan@yonsei.ac.kr

Abstract

In this study, we aimed to define the effects of anthropomorphism and conversational relevance of chatbots on user experience. In specific, the chatbot designed for this study was an online shopping assistant that recommends items for consumers. Levels of anthropomorphism was manipulated by the name, profile picture, word choices, and emojis, while conversational relevance was adjusted by the depth and accuracy of the recommendation. Three categories of user experience were measured: psychological distance, usability, and purchase intentions. The results implied a significant main effect of conversational relevance on all variables for the high anthropomorphized conditions, while all but psychological distance was significant for low anthropomorphized conditions. Although there was no significant main effect of anthropomorphism observed for the variables, the main effect of anthropomorphism on responsibility was marginally significant for a specific item. The results of this study may function as a guidance for future studies regarding usage of chatbots within a marketing setting.

Keywords: Chatbot, Anthropomorphism, CASA, Conversational Relevance, Psychological Distance

1. INTRODUCTION

Increasing numbers of marketing industries are starting to implement the use of ‘chatbots’ as a means of providing service for consumers [1]. Chatbots are known to be an advantage for both the marketers and the users due to its economic efficiency, increased levels of consumer satisfaction, and greater accessibility [2]. However, there still exists shortcomings such as poor levels of comprehension, unnatural flow of conversation, and recurring errors which overturns what may have been a pleasant experience with a chatbot to a frustrating one [2].

The effectiveness of chatbots within the realm of marketing may be explained by the Computers Are Social Actors (CASA) theory. The CASA theory posits that when a computer exudes a social characteristic, the human users tend to project heuristics that have been predefined from human-to-human interactions to human-computer interactions as well [3]. Since interactions with chatbots are highly like those that humans engage with each other within an online platform (i.e., Facebook Messenger, KakaoTalk), it is natural for users to show a tendency to regard chatbots as a social actor, much like a fellow human being [4].

This study integrates the CASA theory with the concept of psychological distance. Psychological distance is derived from the construal level theory, which states that one tends to use highly detailed words that are high in construal to describe actors and items that are considered psychologically intimate. On the other hand, one tends to describe actors and items that are psychologically distant by using words that are low in construal. This dynamic remains true the other way around; when an actor uses a high construal description within a conversation, one tends to regard the actor as being psychologically intimate, and vice versa [5]. In this study, construal level is operationalized as conversational relevance [6].

One of the main qualities of amplifying the social identity of computers is its “humanness”, which is conceptualized as anthropomorphism. Anthropomorphism is the attribution of humanlike characteristics to nonhuman actors—in this case, chatbots [7]. Anthropomorphism is often put into use to increase the level of perceived humanness of artificial intelligence [8]. The level of anthropomorphism may be manipulated by giving the AI agent a humanlike name (i.e., Siri, Genie, Alexa), by modifying the hardware design as such that it resembles humanlike features, and by using a human voice in case of voice-user interfaces [9]. This indicates that as the level of anthropomorphism increases, the perceived level of psychological intimacy increases as well [9].

The current study aims to discover consumer experience of chatbots, specifically focusing on the effects of anthropomorphism and different levels of construal, defined by conversational relevance.

2. THEORY

2.1 Computers Are Social Actors (CASA) Theory

The CASA theory is built upon the assumption that when a computer exhibits characteristics of a social actor, human users tend to interact with it as they would with another human being [3]. There are numbers of research that have attempted to integrate the CASA theory with various fields of psychology. For instance, there is a relationship between display characteristics and social evaluation, whereby the users tend to evaluate their user experience according to the social characteristic such as information transparency and mitigation of anxiety, that are exhibited by the display [10]. Also, a human interaction study regarding praise and blame feedback was replicated on a human-computer interaction and reaped equal results, with a slight difference in the depth of psychological processing on the human’s end [11]. Likewise, the CASA theory functions as a vital foundation for the current article, allowing the premise that users of conversational agents will perceive the system as a social being, therefore interacting with it as they would with a human being.

2.2 Psychological Distance and Construal Level Theory

Psychological distance is defined as the spatial, temporal, social, or hypothetical distance conceived between a stimulus and the actor [12, 13]. The concept of psychological distance is applicable to multiple domains of psychological research. For instance, we tend to report decreased psychological distance in face of events primed with higher levels of emotional intensity [14]. There have been studies integrating psychological distance with the level of perceived feasibility of an object as well. Psychological distance also differed by the perceived feasibility of probability outcomes, where users reported decreased levels of psychological distance to outcomes that are “not as desirable but not as feasible” as compared to those that are “highly desirable but not as feasible” [15]. To add, results from another study revealed that participants that received a “feasible” gift tend to consider the giver as psychologically intimate as compared to those that received a “desirable” gift [16].

Psychological distance is often explained in the context of construal level theory, which posits that the way actors perceive a stimulus depends on the psychological distance with the stimulus [12]. Construal levels are dichotomized into high-level construal and low-level construal, where high-level construal is an abstract, general psychological depiction of a stimulus while low-level construal is a detailed, specific psychological depiction of a stimulus [5]. When the psychological distance between the actor and the stimulus increases, the actors show a tendency to use high-level construal in interpreting the characteristics of the given stimuli. On the other hand, when the psychological distance decreases between the actor and the stimulus, a low-level construal is used to define the stimulus [5]. For instance, the results of a study integrating politeness, psychological distance, and construal levels revealed a higher tendency for the participants to be polite when they thought of the agents as psychologically distant, by using high-level construal words [17]. The opposite causality was held true as well – participants exhibited politeness in front of subjects they construed on a higher level, subsequently reporting the subjects as psychologically distant [17]. Likewise, the concept of psychological distance and construal level theory are mutually complementary, thereby bidirectional.

2.3 Anthropomorphism

Anthropomorphism is defined as the attribution of humanlike characteristics to nonhuman actors — in this case, artificial agents [7]. Of the numerous tools that improve the usability of artificial agents, anthropomorphism is specifically used to heighten the perceived humanness of the agents [8]. Determinants such as the knowledgeable-ness of the artificial agents, the user's personality traits and openness to novel theories, and cultural backgrounds affect the humanness of the artificial agents [18]. For instance, the user's need for human interaction whilst interacting with a chatbot moderates the relationship between the anthropomorphized chatbot and the user's willingness to adapt to the errors made by the chatbots, suggesting that anthropomorphism may possibly be used as a potential solution to bettering human-computer interactions [19].

Anthropomorphizing artificial agents generates multitude of benefits, including predictability, familiarity, adaptability, and accessibility [20]. For example, in cases of dealing with artificial agents that require usage training, first-time users that interact with highly anthropomorphized agents exhibited improved levels of training efficiency [20]. In the context of marketing and consumption, customers showed preference towards anthropomorphized artificial service agents since they were able to readily apply the social cues and heuristics that have been previously used in a human-human interaction [21]. It is also notable that the use of certain communicative behaviors which increase the conversational agent's humanness contributes to a positive outcome in terms of chatbot usability such as intention to use and the sense of cooperation [22]. This implies that an appropriate implementation of anthropomorphism when designing for conversational agents may result in better user experience. By this, the present research grounds on the CASA theory and aims to define how customer's perception of psychological distance, agent usability and purchase intention of a product are affected by the level of anthropomorphism of the service agent.

A chatbot's social characteristic may be categorized into three groups: conversational intelligence, social intelligence, and personification [23]. The current study controlled for anthropomorphism by modifying the chatbot's social characteristics that are known to affect the level of humanness within artificial agents. Based on previous literature, traits such as the artificial agent's name, profile picture, and speech patterns are some of the most effect ways to alter the level of anthropomorphism perceived by the users. For instance, implementing human names and language styles to artificial agents heightened the rate of which the users of the agent perceived it to be more humanlike [24]. In specific, a socially oriented conversation style of the artificial agent enhanced the user's perception of the chatbot being a social entity [25]. To add, a set of meta-

analyses implied a consistent pattern within anthropomorphism studies, whereby a robot's physical embodiment tends to increase the level of humanness [21]. Consistent findings as these allow the current study to use physical features of the conversational agents as one of the key elements to manipulating the levels of perceived humanness of chatbots.

In addition, this study implements the use of 'emoji' to control for the level of humanness of conversational agents. There have been findings indicating that the use of emoji in chatbot conversation increases the level of social attraction as well as the competence and credibility of the artificial agent [4]. Previous research indicates that the use of emoji as a marketing scheme may attribute a reliable and significant personality to brands [26]. To add, emoji usage within chatbot conversations may increase the humanness of the agent due to the heavily contextual nature of the emoji itself [27]. Likewise, past studies on emoji implies that appropriate usage may in fact heighten the perceived humanness of chatbots by ascribing a personality to the agent as well as implying that the chatbot is indeed capable of making extensive conversation in depth. However, there has been less work done regarding the dynamic between emoji and its effect on anthropomorphism. By this, the current study aims to close this knowledge gap by manipulating emoji usage as one of the key elements of humanness in chatbots, thereby examining how the use of emojis impact the perceived level of anthropomorphism.

2.4 Conversational Relevance

Conversational relevance is determined by the relatedness of an agent's response to an actor [28]. High conversational relevance is established when the agents consider the background information of the discussion at hand and conveys the underlying intention of the actor [28]. When a human agent's response to an actor's input is highly related to the matter at hand in a human-human interaction, the interaction is built upon high conversational relevance [6]. On the other hand, when a human agent's response to an actor's input is unrelated to the question of interest, the interaction is considered to have low conversational relevance [6]. An equal dynamic holds true to human-computer interactions, in that when the artificial agent's output of a user's input is highly related to the context, the interaction is rated as having high conversational relevance; when the agent's output is unrelated to the interactional context, the conversational relevance is low [6].

Existing research shows that highly anthropomorphized artificial agents trigger high expectations of the interaction that the agent will provide [29]. However, when this expectation is unmet, the user's attitude towards the artificial agent will plummet, including the attitude towards the artificial agent, usability, and likeability [29]. Therefore, it is vital to align the levels of anthropomorphism and the user service the agent can provide. The current study operationalizes construal levels of the artificial agent's output as conversational relevance. This mechanism is based on a previous observation of the relationship between conversational relevance and construal levels, whereby higher-level construal (defined as shared-knowledge information) has a strong effect on increasing conversational relevance of an interaction, rendering the two concepts as interchangeable [30]. Due to the high expectations users hold against highly anthropomorphized artificial agents [29], it is expected that the expectancy violation scores will be higher when a highly anthropomorphized conversational agent provides services with low conversational relevance, as compared to when a low anthropomorphized conversational agent provides services with low conversational relevance, since the expectations for low anthropomorphized conversational agents would be lower to begin with.

2.5 Chatbot in E-Commerce

With the increasing consumption of online commerce, companies and developers are striving to devise a unique and effective method to communicate with their potential customers, primarily by making use of online

communicative platforms such as social networking services and artificial agents. There are increasing instances of the use of chatbots in terms of customer services, as well as research aimed towards bettering the user experience of chatbots. Sufficient use of chatbots in marketing can bring about great benefits, such as eliciting brand familiarity as well as being able to emphasize the utilitarian value of a product [31]. One of the crucial elements that are known to heighten the overall usability of chatbots is the perceived humanness of the agent, in other words, the level of anthropomorphism. There have been findings which suggest that anthropomorphism have a positive influence on consumer’s trust in chatbots, which subsequently heightens the purchase intention of the product the chatbot is recommending to the user [32]. As the perception of the chatbot’s anthropomorphism increased, users reported higher levels of social presence as well as enjoyment, which then led to increased purchase intentions regarding the item the chatbot is recommending [33]. By this, the current study additionally tests for the effects of anthropomorphism of conversational agents on the user’s purchase intentions, specifically designed to aid the customer’s shopping through product recommendation.

As shown in Figure 1, the hypotheses of this study are as follow:

H1: Expectations regarding human-computer interactions will be higher when presented with a high anthropomorphized chatbots as compared to low anthropomorphized chatbots.

H2: When both the level of anthropomorphism and a) conversational relevance of the chatbot are high, the users will rate highly of the 1) psychological intimacy, 2) usability, and 3) purchase intentions b) as compared to the equally anthropomorphized chatbot with low level of conversational relevance.

H3: When the level of anthropomorphism is low and a) the conversational relevance of the chatbot is high, the users will rate highly of the 1) psychological intimacy, 2) usability, and 3) purchase intentions b) as compared to the equally anthropomorphized chatbot with low level of conversational relevance.

H4: There will be an interaction effect of anthropomorphism and conversational relevance in terms of 1) psychological intimacy, 2) usability, and 3) purchase intentions.

H5: Regardless of the level of conversational relevance, the measures for all variables will be higher for high anthropomorphized chatbots as compared to low anthropomorphized chatbots.

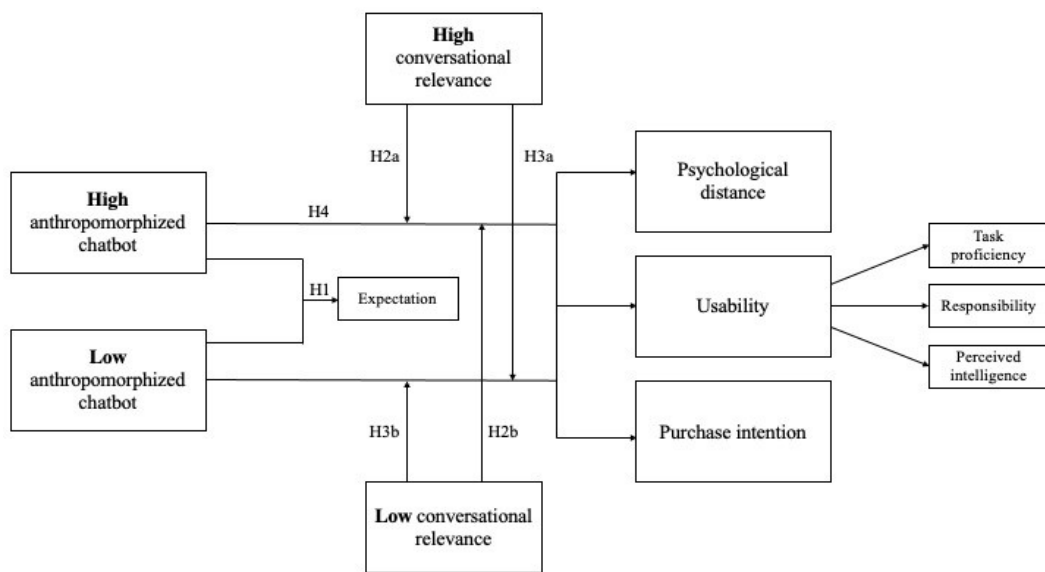


Figure 1. Conceptual framework and hypotheses

3. METHODS

3.1 Participants

A 2 (anthropomorphism: high vs. low) \times 2 (conversational relevance: high vs. low) between-subjects design was used to test the hypotheses. A total of 120 Yonsei University undergraduates comprising of 62.5% females and 37.5% males, with the mean age of 21.13 years ($SD = 3.53$) were recruited through the Yonsei Psychology Research Participation System (Yonsei SONA System) and were randomly assigned to one of the four conditions.

3.2 Materials

The entire process was conducted online through Qualtrics, and one credit was given to every participant who completed the survey by the deadline. The experiment lasted for approximately 5 to 10 minutes. All chatbot conversations were fabricated by chat screenshot generator [34]. The Facebook Messenger interface was used to control for familiarity biases that may occur due to differences in message application interfaces based on the type of smartphones (see Appendix A). High anthropomorphized chatbot conversations consisted of 1) a humanlike name (“Mina”), 2) profile picture of a humanlike avatar, 3) colloquial speech patterns, and 4) emoji usage. On the other hand, low anthropomorphized chatbot conversations consisted of 1) a non-humanlike name (“Shopbot”), 2) profile picture of a non-humanlike icon, 3) literate speech patterns, and 4) no emoji usage.

3.3 Procedures

First, a sample chatbot conversation consisting of a 1) single input from a virtual user and 2) a reply of self-introduction from the chatbot which differed across anthropomorphism conditions (see Figure 1) was given to the participants. The participants were asked to take time to thoroughly digest the sample conversations. Afterwards, a series of surveys were given for the participants to fill out (see Appendix B). The participants were immediately asked to complete a short survey measuring the level of 1) anthropomorphism and 2) expectancy based on the sample conversations. The anthropomorphism scale was based on “Godspeed I: Anthropomorphism”, a survey consisting of five items developed to aid developers to measure the level of humanness of robots and artificial intelligence in a concise and precise manner [35]. The results from these five items measuring anthropomorphism were later used to check whether the anthropomorphism manipulation was successful. The expectancy scale was given as a single item, based on an expectancy manipulation check questionnaire [8]. Afterwards, the participants were given three sets of conversation screenshots consisting of a virtual interaction of a user and a shopping assistant chatbot, with the overall flow of the conversation being similar except for 1) the suggested item differing across the sets (headphones, sunglasses, pen), as well as 2) the additional closing lines coming from the anthropomorphized chatbot, to increase humanness through natural, colloquial speech patterns. The items selected for suggestion were based on social distance and gift preference, where they confirmed that pen, sunglasses, headphones, leather gloves, and wallet to be representative of material goods as compared to experiential goods [36]. Since past literature indicated that purchasing of material goods required less information on personal traits as compared to experiential goods, it was considered appropriate to use material goods in the current study due to the non-intrusive nature and the brevity of the experiment [36]. After each conversation screenshot, the participants were asked to fill out a survey measuring the level of 1) conversational relevance, a single item (“The chatbot provided the user with an appropriate response”), which was used as a manipulation check, 2) three items of psychological distance [9], 3) seven items of perceived usability [35], and 4) five items of purchase intentions [37]. All survey items

were modified to a 7-point Likert scale in order to maintain consistency. A demographic survey including the age and gender of the participant was administered at the very end of the experiment.

4. RESULTS

4.1 Manipulation Check

A two-way analysis of variance (ANOVA) was used to test the hypotheses of the experiment. The main effect of the level of anthropomorphism was significant, $t(118) = 2.43, p = 0.02$, with participants in the high anthropomorphized condition scoring higher on the anthropomorphism scale ($M = 14.35, SD = 3.81$) compared to those in the low anthropomorphized condition ($M = 12.65, SD = 3.84$), indicating that the manipulation of anthropomorphism was successful. The main effect of conversational relevance was significant as well, $t(118) = 14.13, p < 0.01$, with participants in the relevant condition scoring higher on the relevance scale ($M = 15.98, SD = 2.70$) compared to those in the irrelevant condition ($M = 7.93, SD = 3.49$), indicating that the manipulation of conversational relevance was successful.

4.2 Conversation Expectation Violation

There was no significant difference between the conversational expectation between the high and low anthropomorphized condition, $t(118) = -0.307, p = 0.76$, with those in the high anthropomorphized condition scoring similarly on the expectation scale ($M = 10.20, SD = 1.98$), as the low anthropomorphized condition ($M = 10.32, SD = 2.18$).

Table 1. Mean and standard deviation for conversation expectancy violation

		Anthropomorphism	
		High	Low
Conversational Relevance	High	5.87 (0.73)	5.33 (1.18)
	Low	3.00 (1.68)	3.17 (1.72)

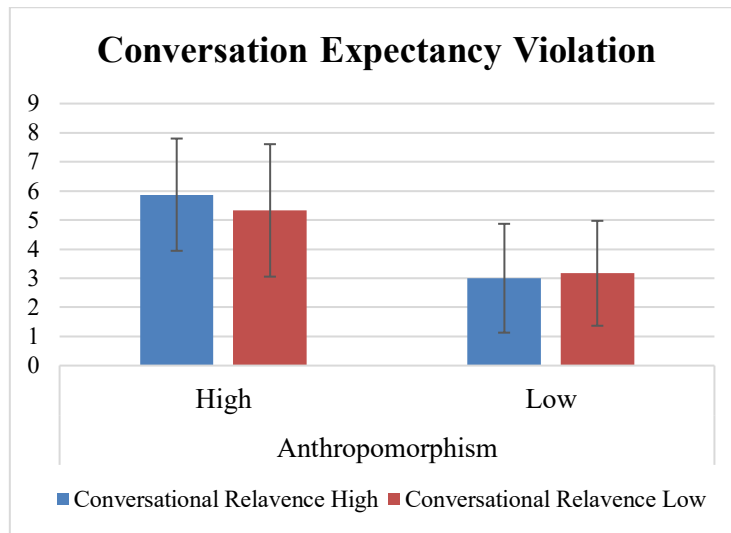


Figure 2. Mean and standard deviation for conversation expectancy violation

There was a significant main effect of conversational relevance on conversation expectancy violation, $F(1, 116) = 98.27, p < 0.01$, with those in the high conversational relevance conditions reporting that the chatbot met their conversation expectations ($M = 5.06, SD = 1.01$), while those in the low conversational relevance conditions reporting that the chatbot failed to meet their conversation expectations ($M = 3.08, SD = 1.69$). However, as shown in Table 1 and Figure 2, there was no significant main effect of anthropomorphism on conversation expectancy violation, $F(1, 116) = 0.52, p = 0.47$, thereby indicating that the difference of conversation expectancy violation scores between high and low conversational relevance conditions within high anthropomorphized conditions ($M = 5.87, SD = 0.73; M = 3.00, SD = 1.68$) and low anthropomorphized conditions ($M = 5.33, SD = 1.18; M = 3.17, SD = 1.72$) was not significant.

4.3 Psychological Distance

As shown in Table 2 and Figure 3, there was a marginal main effect of conversational relevance on psychological distance within the high anthropomorphized condition, $F(1, 58) = 3.87, p = 0.05$, with those in the high conversational relevance condition reporting marginally heightened level of psychological intimacy ($M = 33.43, SD = 10.56$) as compared to those in the low conversational relevance condition ($M = 27.57, SD = 12.46$). On the other hand, there was no significant main effect of conversational relevance observed within the low anthropomorphized condition, $F(1, 58) = 0.04, p = 0.85$, with those in the high conversational relevance condition reporting similar levels of psychological intimacy ($M = 22.33, SD = 10.24$) as those in the low conversational relevance condition ($M = 21.83, SD = 9.87$).

Table 2. Mean and standard deviation for psychological distance

		Anthropomorphism	
		High	Low
Conversational Relevance	High	33.43 (10.56)	22.33 (10.24)
	Low	27.57 (12.46)	21.83 (9.87)

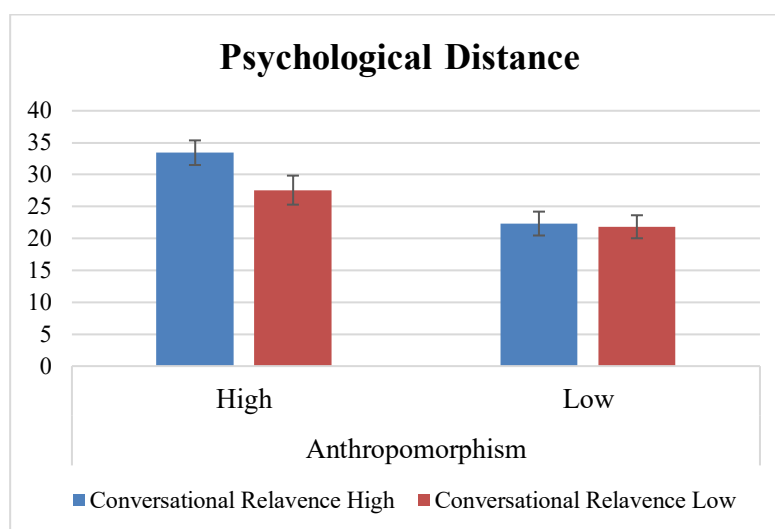


Figure 3. Mean and standard deviation for psychological distance

However, there was a significant main effect of anthropomorphism observed for psychological distance, with the numbers for high ($M = 30.50, SD = 11.83$) and low ($M = 22.08, SD = 9.97$) anthropomorphized conditions, $F(1, 118) = 17.75, p < 0.01$, being presented as such. These main effects were not qualified by an interaction between anthropomorphism and conversational relevance, $F(1, 116) = 1.84, p = 0.18$.

4.4 Usability

Three items were measured to predict the usability of chatbots: task proficiency, perceived intelligence, and perceived responsibility.

4.5 Task Proficiency

As shown in Table 3 and Figure 4, There was a significant main effect of conversational relevance on task proficiency within the high anthropomorphized conditions, $F(1, 58) = 56.58, p < 0.01$, with those in the high conversational relevance condition reporting higher scores for the chatbot’s task proficiency ($M = 16.30, SD = 2.28$) as compared to those in the low conversational relevance condition ($M = 10.00, SD = 3.98$). There was a significant main effect of conversational relevance on task proficiency within the low anthropomorphized conditions as well, $F(1, 58) = 66.77, p < 0.01$, with those in the high conversational relevance condition reporting higher scores for task proficiency ($M = 15.93, SD = 2.38$) as compared to those in the low conversational relevance condition ($M = 8.90, SD = 4.07$).

Table 3. Mean and standard deviation for task proficiency

		Anthropomorphism	
		High	Low
Conversational Relevance	High	16.30 (2.28)	15.93 (1.38)
	Low	10.00 (3.98)	8.90 (4.07)

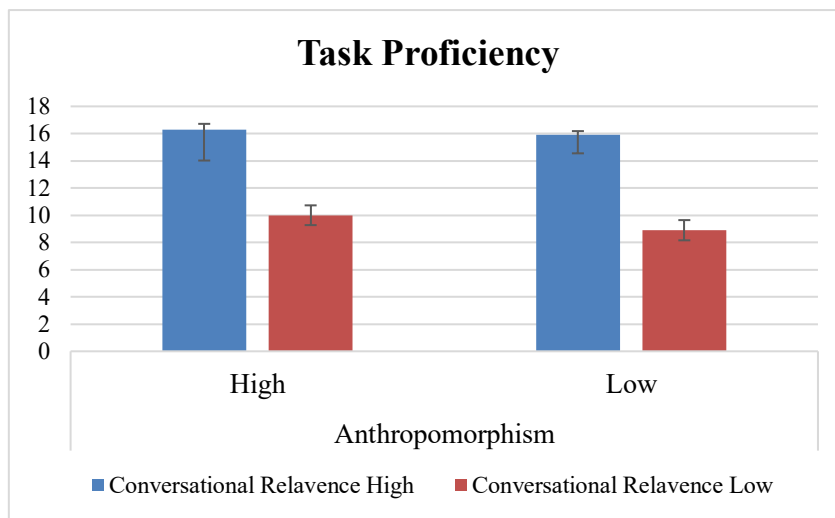


Figure 4. Mean and standard deviation for task proficiency

However, there was no significant main effect of anthropomorphism observed for the scores for task proficiency, with the numbers for the high ($M = 13.15$, $SD = 4.52$) and low ($M = 12.42$, $SD = 4.85$) anthropomorphized conditions, $F(1, 118) = 0.73$, $p = 0.40$, being presented as such. These main effects were not qualified by an interaction between anthropomorphism and conversational relevance, $F(1, 116) = 0.37$, $p = 0.54$.

4.6 Perceived intelligence

As for perceived intelligence, there was a significant main effect of conversational relevance observed within the high anthropomorphized condition, $F(1, 58) = 24.30$, $p < 0.01$. As shown in Table 4 and Figure 5, participants in the high conversational relevance condition perceiving higher levels of intelligence for the chatbot ($M = 40.63$, $SD = 10.91$) as compared to those in the low conversational relevance condition ($M = 26.83$, $SD = 10.77$). There was a significant main effect of conversational relevance on perceived intelligence within the low anthropomorphized conditions as well, $F(1, 58) = 26.17$, $p < 0.01$, with those in the high conversational relevance condition perceiving higher levels of intelligence for the chatbot ($M = 38.30$, $SD = 10.16$) as compared to those in the low conversational relevance condition ($M = 25.50$, $SD = 9.19$).

Table 4. Mean and standard deviation for perceived intelligence

		Anthropomorphism	
		High	Low
Conversational Relevance	High	40.63 (10.91)	38.30 (10.16)
	Low	26.83 (10.77)	25.50 (9.19)

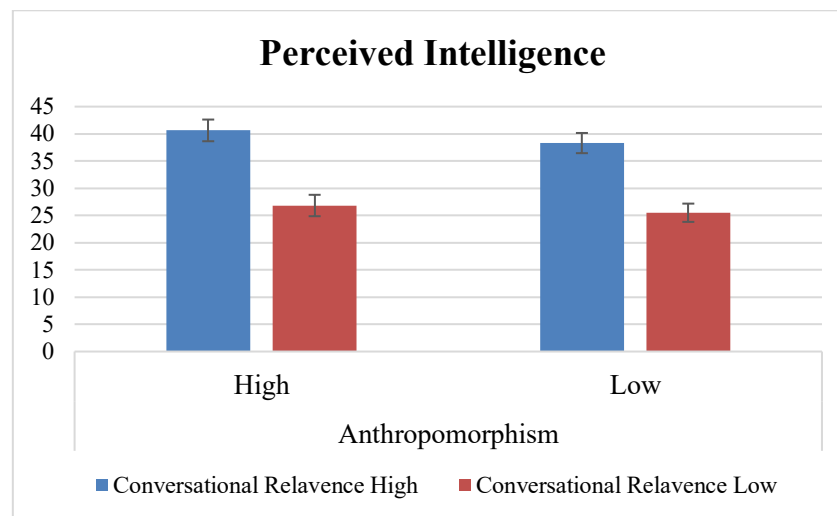


Figure 5. Mean and standard deviation for perceived intelligence

However, there was no significant main effect of anthropomorphism observed for the perceived level of intelligence, with the numbers for the high ($M = 33.73$, $SD = 12.84$) and low ($M = 31.90$, $SD = 11.57$) anthropomorphized conditions, $F(1, 118) = 0.68$, $p = 0.41$, being as such. These main effects were not qualified by an interaction between anthropomorphism and conversational relevance, $F(1, 116) = 0.07$, $p = 0.79$.

4.7 Responsibility

The last item measured for usability was the perceived level of responsibility. As shown in Table 5 and Figure 6, there was a significant main effect of conversational relevance on perceived responsibility within the high anthropomorphized conditions, $F(1, 58) = 5.65, p = 0.02$, with the ratings for perceived responsibility of the chatbot being higher for participants in the high conversational relevance condition ($M = 13.97, SD = 4.19$) as compared to those in the low conversational relevance condition ($M = 11.27, SD = 4.60$). There was a significant main effect of conversational relevance on perceived responsibility within the low anthropomorphized conditions as well, $F(1, 58) = 9.77, p < 0.01$, with the ratings for perceived responsibility of the chatbot being higher for participants in the high conversational relevance condition ($M = 12.73, SD = 3.66$) as compared to those in the low conversational relevance condition ($M = 9.73, SD = 3.76$).

Table 5. Mean and standard deviation for responsibility

		Anthropomorphism	
		High	Low
Conversational Relevance	High	13.97 (4.19)	12.73 (3.66)
	Low	11.27 (4.60)	9.73 (3.76)

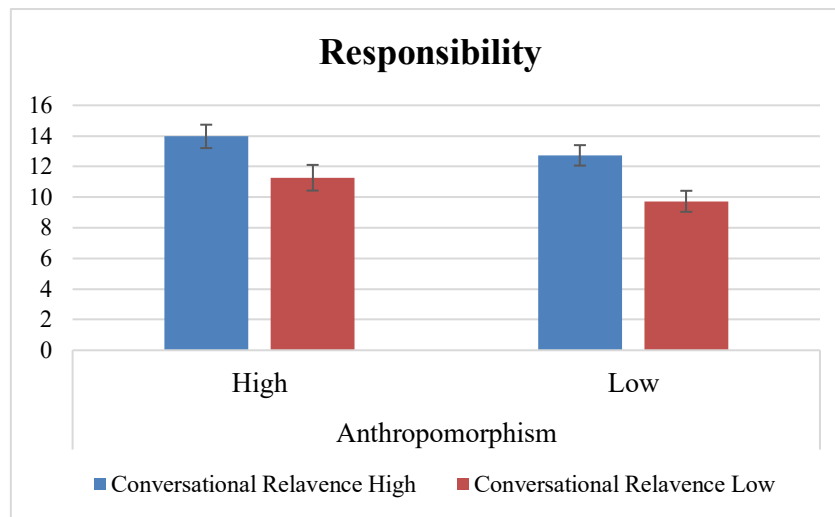


Figure 6. Mean and standard deviation for responsibility

However, there was no significant main effect of anthropomorphism observed for the perceived level of responsibility, with the numbers for the high ($M = 12.62, SD = 4.57$) and low ($M = 11.23, SD = 3.98$) anthropomorphized conditions, $F(1, 118) = 3.12, p = 0.08$, being as such. One thing to note was a significant main effect of anthropomorphism on responsibility observed for the conversation screenshot in which the chatbot recommended headphones for the user, $F(1, 118) = 4.09, p = 0.05$, with participants in the high anthropomorphized conditions giving higher ratings for the perceived responsibility of the chatbot that recommended headphones ($M = 4.42, SD = 1.63$) compared to those in the low anthropomorphized conditions ($M = 11.23, SD = 3.98$). These main effects were not qualified by an interaction between anthropomorphism and conversational relevance, $F(1, 116) = 0.04, p = 0.84$.

4.8 Purchase intention

As shown in Table 6 and Figure 7, here was a significant main effect of conversational relevance observed for purchase intention within the high anthropomorphized conditions, $F(1, 58) = 28.74, p < 0.01$, with purchase intention for items recommended by the chatbot being higher for participants in the high conversational relevance condition ($M = 55.80, SD = 14.67$) as compared to those in the low conversational relevance condition ($M = 35.23, SD = 15.04$). Likewise, there was a significant main effect of conversational relevance on purchase intention within the low anthropomorphized conditions as well, $F(1, 58) = 64.91, p < 0.01$, with the purchase intention for those in the high conversational relevance condition being higher ($M = 56.50, SD = 12.17$) as compared to those in the low conversational relevance condition ($M = 30.50, SD = 12.82$).

Table 6. Mean and standard deviation for purchase intention

		Anthropomorphism	
		High	Low
Conversational Relevance	High	55.80 (14.67)	56.50 (12.17)
	Low	35.23 (15.04)	30.50 (12.82)

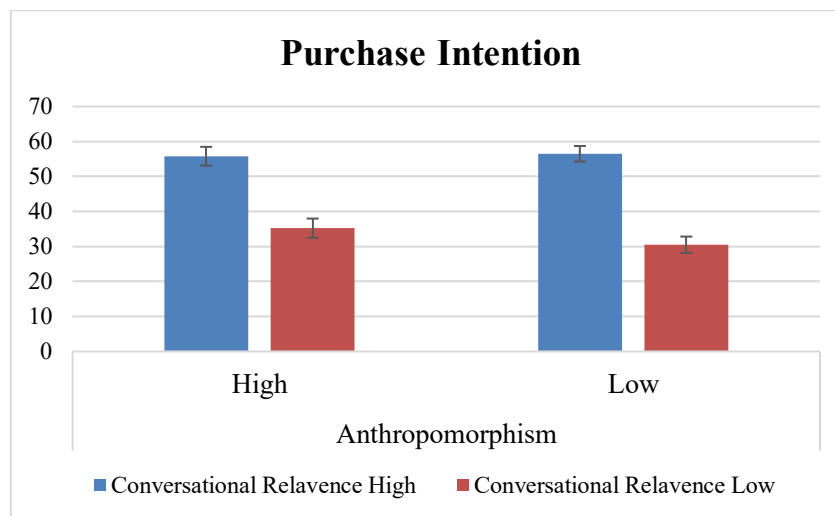


Figure 7. Mean and standard deviation for purchase intention

However, there was no significant main effect of anthropomorphism observed for purchase intention, with the numbers for high ($M = 45.51, SD = 18.02$) and low ($M = 43.50, SD = 18.04$) anthropomorphized conditions, $F(1, 118) = 0.38, p = 0.54$, being as such. These main effects were not qualified by an interaction between anthropomorphism and conversational relevance, $F(1, 116) = 1.18, p = 0.28$.

5. DISCUSSION

This study aimed to define the effects of anthropomorphism and construal level on a conversational agent, and how qualities as such impacts the perceived levels of psychological distance, the usability of the chatbot, and the purchase intention of the items recommended by the chatbot. The manipulation of the level of anthropomorphism and construal was successful; however, contrary to H1, the expectations between the different levels of anthropomorphized chatbots were not significant. The statistics imply that the overall

expectations for the chatbot interaction was low across all conditions, which may indicate that the participants may have a low preference for chatbots in general. Although the current study used emojis on the high anthropomorphized condition based on past research regarding heightened level of credibility on chatbots with emoji use [4], it seems that implementing emojis were insufficient to cause significant difference in conversational expectation.

There was no significant interaction effect observed in the current experiment, thereby rejecting H4. This contradiction may be due to several different causes. First, the chatbot's social characteristics may be classified into three categories: conversational intelligence, social intelligence, and personification [23]. Anthropomorphism, in other words, the chatbot's "humanness" is categorized as an element of social intelligence, while conversational relevance falls into the category of conversational intelligence [23]. Although both categories represent the chatbot's social characteristics, the subtle difference that distinguishes conversational and social intelligence may have influenced the minimal results in terms of interaction effects. By extension, while questionnaires designed to evaluating the psychological distance of chatbots were targeted towards evaluating the social intelligence of the chatbot, the other two grand categories (usability and purchase intention) were aimed towards gauging conversational intelligence. To add, there was an additional closing line added in the high anthropomorphized conversation in order to intensify the humane essence of natural speech. However, subtle inconsistencies as such may have influenced the outcome of the overall experiment.

Second, there may have been an overall lack of credibility and trustworthiness towards the chatbot that have been used in the experiment. Some of the crucial factors that affect the perceived credibility of service chatbots include the quality of output given by the chatbot, perceived security and privacy, and the brand value of the platform which the chatbot is hosted on [38]. By this, developing the chatbot used in the present research by implementing indication of security and privacy as well as specifying the hosting brand of the chatbot may lead to different yet meaningful results.

Third, the chatbot used in the current study may have fallen short in exhibiting sufficient social presence. There has been research suggesting that increased social presence of a chatbot enhances the overall positivity of the interaction [25]. Affective characteristics and relational cues are some of the decisive elements that affect the social presence of a chatbot [25]. In future research, providing the participants with a chance to actively interact with the chatbot rather than presenting them with a screenshot of a sample interaction may lead to an increase in the social presence of the chatbot, thereby leading to different results than the current study.

Moreover, latent variables of the participants such as personality traits and their openness to novel technology may have affected the results of the study. Technology acceptance model (TAM) refers to the process which the user's reaction to information technology leads to or does not lead to intentions to use information technology, which then may or may not lead to actual usage [39]. TAM is known to be mainly affected by two factors – perceived usefulness and perceived ease of use [40]. The intensity of these factors may differ by the user's experience with technology, socioeconomic and cultural background, and even personality. These factors may function as a means to heighten user experience in the context of artificial agent usage by implementing deep learning methods, like those that are being used in SNS programming [41]. Therefore, in future research, it may be interesting to measure latent variables as such to reveal the dynamics that have reaped such results.

The main effects of conversational relevance on usability and purchase intentions were significant across both high and low anthropomorphized conditions. One thing to note was that although the main effects of conversational relevance on psychological distance was significant regarding conversational relevance within high anthropomorphized conditions, the effect was not significant in the case for low anthropomorphized

conditions. This implies a partial support for H3 while there was full support for H2. A possible interpretation would be that since the perceived psychological distance of the low anthropomorphized chatbot was comparably low as compared to the high anthropomorphized chatbot, the effect of high conversational relevance was insufficient to decrease the level of psychological distance.

Lastly, there was a significant main effect of anthropomorphism on psychological distance, while there was no significant main effect for usability nor purchase intention, showing partial support for H5. However, the responsibility factor within the usability scale was marginally significant; specifically, there was a significant main effect of anthropomorphism on responsibility observed for chatbot screenshots recommending headphones. This result is intriguing that there was a difference in the rating of the variables depending on the nature of the recommended item. For instance, experiential goods such as a concert ticket and a travel package are considered as requiring very specific knowledge of the buyer or the recipient, whereas material goods such as the ones used in the current study are known to require less of the personal characteristics of the buyer or the recipient, but rather a safe choice that are inclusive of the preference of the general population [36]. The current study failed to diversify the types of goods that have been suggested by the chatbot, since all three items (headphones, glasses, and pen) were strictly material and have not taken into consideration the personal likings of the participants. Likewise, future studies may be designed to reveal the dynamics between perceived responsibility and the characteristics of the items recommended by the chatbot, by categorizing different types of goods that may be offered by the artificial agent.

6. CONCLUSION

Although the future and development of artificial agents are of great interest to many, the firsthand users of the technology continue to report the various limitations and shortcomings they are often faced with during agent interaction. The level of trust given to the artificial agents, including voice user interfaces (ex. Siri, Bixby, Alexa), chatbots (ex. chatbots in KakaoTalk, Facebook), and robots, are dwindling. This study was focused on revealing the means to increasing the overall user experience of chatbots, by controlling for the level of anthropomorphism, in other words, the ‘humanness’ of chatbots, and conversational relevance. The result implied that the overall trust of the system itself is lacking, regardless of the level of humanness of the chatbot. However, the users tend to be very sensitive about the conversational relevance of the chatbot’s dialogues, implying that the key to increasing usability of an artificial agent may lie in the content and the quality of the agent’s response. Future research may focus on elaborating on these results, by testing specific factors that may affect the overall quality and sophistication of the agent’s response.

ACKNOWLEDGEMENT

This work was supported by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NRF- 2019S1A5C2A03083499).

REFERENCES

- [1] E. Van den Broeck, B. Zarouali, and K. Poels, “Chatbot advertising effectiveness: When does the message get through?” *Computers in Human Behavior*, Vol. 98, pp. 150-157, September 2019, <https://doi.org/10.1016/j.chb.2019.04.009>
- [2] C. Popescu, “Chatbots as marketing communication tool,” *FAIMA Business & Management Journal*, Vol. 8, No. 3, pp. 62-75, September 2020.
- [3] B. Reeves and C. Nass, *The media equation: How people treat computers, television, and new media like*

real people and places. Center for the Study of Language and Information, Cambridge University Press, 1996.

- [4] A. Beattie, A. P. Edwards, and C. Edwards, "A bot and a smile: Interpersonal impressions of chatbots and humans using emoji in computer-mediated communication," *Communication Studies*, Vol. 71, No. 3, pp. 409-427, February 2020, <https://doi.org/10.1080/10510974.2020.1725082>
- [5] Y. Trope and N. Liberman, "Construal-level theory of psychological distance," *Psychological Review*, Vol. 117, No. 2, pp. 440-463, April 2020, doi: 10.1037/a0018963
- [6] R. M. Schuetzler, J. S. Giboney, G. M. Grimes, and J. F. Nunamaker, J. F., "The influence of conversational agent embodiment and conversational relevance on socially desirable responding," *Decision Support Systems*, Vol. 114, pp. 94-102, August 2018, <https://doi.org/10.1016/j.dss.2018.08.011>
- [7] American Psychological Association, *Anthropomorphism*. <https://apastyle.apa.org/style-grammar-guidelines/grammar/anthropomorphism>.
- [8] G. M. Grimes, R. M. Schuetzler, and J. S. Giboney, "Mental models and expectation violations in conversational AI interactions," *Decision Support Systems*, Vol. 144, pp. 1-10, January 2021, <https://doi.org/10.1016/j.dss.2021.113515>
- [9] X. Li and Y. Sung, "Anthropomorphism brings us closer: The mediating role of psychological distance in User-AI assistant interactions," *Computers in Human Behavior*, Vol. 118, pp. 1-9, January 2021, <https://doi.org/10.1016/j.chb.2021.106680>
- [10] E., Robles, C. Nass, and A. Kahn, "The social life of information displays: How screens shape psychological responses in social contexts," *Human-Computer Interaction*, Vol. 24, pp. 48-78, April 2009, DOI: 10.1080/07370020902739320
- [11] P. Mishra, "Affective feedback from computers and its effect on perceived ability and affect: A test of the computers as social actor hypothesis," *Journal of Educational Multimedia and Hypermedia*, Vol. 15, No. 1, pp. 107-131, January 2006.
- [12] Y. Bar-Anan, N. Liberman, and Y. Trope, "The association between psychological distance and construal level evidence from an implicit association test," *Journal of Experimental Psychology: General*, Vol. 135, No. 4, pp. 609-622, November 2006, DOI: 10.1037/0096-3445.135.4.609
- [13] L. E. Park, A. F. Young, and P. W. Eastwick, "(Psychological) distance makes the heart grow fonder: Effects of psychological distance and relative intelligence on men's attraction to women," *Personality and Social Psychology Bulletin*, Vol. 41, No. 11, pp. 1459-1473, November 2015, DOI: 10.1177/0146167215599749
- [14] L. Van Boven, J. Kane, A. P. McGraw, and J. Dale, "Feeling close: Emotional intensity reduces perceived psychological distance," *Journal of Personality and Social Psychology*, Vol. 98, No. 6, pp. 872-885, June 2010, DOI: 10.1037/a0019262
- [15] A. Todorov, A. Goren, and Y. Trope, "Probability as a psychological distance: Construal and preferences," *Journal of Experimental Social Psychology*, Vol. 43, pp. 473-482, May 2007, <https://doi.org/10.1016/j.jesp.2006.04.002>
- [16] S. Y. Rim, K. E. Min, P. J. Liu, T. L. Chartrand, and Y. Trope, "The gift of psychological closeness: How feasible versus desirable gifts reduce psychological distance to the giver," *Personality and Social Psychology Bulletin*, Vol. 45, No. 3, pp. 360-371, July 2018, <https://doi.org/10.1177/0146167218784899>
- [17] E. Stephan, N. Liberman, and Y. Trope, "Politeness and psychological distance: A construal level perspective," *Journal of Personality and Social Psychology*, Vol. 98, No. 2, pp. 268-280, February 2010, DOI: 10.1037/a0016960
- [18] N. Epley, A. Waytz, and J. T. Cacioppo, "On seeing human: A three-factor theory of anthropomorphism," *Psychological Review*, Vol. 114, No. 4, pp. 864-886, October 2007, DOI: 10.1037/0033-295X.114.4.864
- [19] B. Sheehan, H. S. Jin, and U. Gottlieb, "Customer service chatbots: Anthropomorphism and adoption,"

- Journal of Business Research*, Vol. 115, pp. 14-24, July 2020, <https://doi.org/10.1016/j.jbusres.2020.04.030>
- [20] J. Zlotowski, D. Proudfoot, K. Yogeewaran, and C. Bartneck, "Anthropomorphism: Opportunities and challenges in human-robot interaction," *International Journal of Social Robotics*, Vol. 7, pp. 347-360, June 2015, <https://doi.org/10.1007/s12369-014-0267-6>
- [21] M. Blut, C. Wang, N. V. Wunderlich, and C. Brock, "Understanding anthropomorphism in service provision: A meta-analysis of phsyical robots, chatbots, and other AI," *Journal of the Academy of Marketing Science*, Vol. 49, pp. 632-658, January 2021, <https://doi.org/10.1007/s11747-020-00762-y>
- [22] M. M. E. Van Pinxteren, M. Pluymaekers, and J. G. A. M. Lemmink, "Human-like communication in conversational agents: A literature review and research agenda," *Journal of Service Management*, Vol. 31, No. 2, pp. 203-225, September 2020, DOI 10.1108/JOSM-06-2019-0175
- [23] A. P. Chaves and M. A. Gerosa, "How should my chatbot interact? A survey on social characteristics in human-chatbot interaction design," *International Journal of Human-Computer Interaction*, Vol. 37, No. 8, pp. 729-758, November 2020, <https://doi.org/10.1080/10447318.2020.1841438>
- [24] T. Araujo, "Living up to the chatbot hype: The influence of anthropomorphic design cues and communicative agency framing on conversational agent and company perceptions," *Computers in Human Behavior*, Vol. 85, pp. 183-189, August 2018, <https://doi.org/10.1016/j.chb.2018.03.051>
- [25] R. De Cicco, S. C. Silva, and F. R. Alparone, "Millennials' attitude toward chatbots: An experimental study in a social relationship perspective," *International Journal of Retail & Distribution Management*, Vol. 48, No. 11, pp. 1213-1233, July 2020, DOI:10.1108/ijrdm-12-2019-0406
- [26] S. Moussa, "Measuring brand personality using emoji: Findings from Mokken scaling," *Journal of Brand Management*, Vol. 28, pp. 116-132, March 2021, <https://doi.org/10.1057/s41262-020-00220-8>
- [27] D. Westerman, A. C. Cross, and P. G. Lindmark, "I believe in a thing called bot: Perceptions of the humanness of "chatbots"," *Communication studies*, Vol. 70, No. 3, pp. 295-312, July 2019, <https://doi.org/10.1080/10510974.2018.1557233>
- [28] M. Dascal, "Conversational relevance," *Journal of Pragmatics*, Vol. 1, No. 4, 1977, pp. 309-328.
- [29] E. C. Ling, I. Tussyadiah, A. Tuomi, J. Stienmetz, and A. Ioannou, "Factors influencing users' adoption and use of conversational agents: A systematic review," *Psychology & Marketing*, Vol. 38, No. 7, pp. 1031-1051, April 2021, <https://doi.org/10.1002/mar.21491>
- [30] A. E. Clark and G. R. Semin, "Receivers' expectations for abstract versus concrete construal: Conversational relevance as a determinant of construal level," *Journal of Language and Social Psychology*, Vol. 27, No. 2, pp. 155-167, January 2008, <https://doi.org/10.1177/0261927X07313645>
- [31] L. Lo Presti, G. Maggiore, and V. Marino, "The role of the chatbot on customer purchase intention: Towards digital relational sales," *Italian Journal of Marketing*, pp. 165-188, June 2021, <https://doi.org/10.1007/s43039-021-00029-6>
- [32] C. Yen and M. Chiang, "Trust me, if you can: A study on the factors that influence consumers' purchase intention triggered by chatbots based on brain image evidence and self-reported assessments," *Behaviour & Information Technology*, Vol. 40, No. 11, pp. 1177-1194, March 2020, <https://doi.org/10.1080/0144929X.2020.1743362>
- [33] M. C. Han, "The impact of anthropomorphism on consumers' purchase decision in chatbot commerce," *Journal of Internet Commerce*, Vol. 20, No. 1, pp. 46-65, January 2021, DOI:10.1080/15332861.2020.1863022
- [34] Geekprank. <https://geekprank.com/chat-screenshot/>.
- [35] C. Bartneck, D. Kulić, E. Croft, and S. Zoghbi, "Measurement instruments for the anthropomorphism, animacy, likability, perceived intelligence, and perceived safety of robots," *International Journal of Social Robotics*, Vol. 1, pp. 71-82, January 2009, <https://doi.org/10.1007/s12369-008-0001-3>
- [36] J. K. Goodman and S. Lim, "When consumers prefer to give material gifts instead of experiences: The role

- of social distance,” *Journal of Consumer Research*, Vol. 45, pp. 365-382, August 2018, <https://doi.org/10.1093/jcr/ucy010>
- [37] N., Barber, N., P. Kuo, M. Bishop, and R. Goodman, “Measuring psychographics to assess purchase intention and willingness to pay,” *Journal of Consumer Marketing*, Vol. 29, No. 4, pp. 280-292, June 2012, DOI 10.1108/07363761211237353
- [38] A. Følstad, C. B. Nordheim, and C. A. Bjørkli, “What makes users trust a chatbot for customer service? An exploratory interview study,” *Internet Science*, pp. 194-208, September 2018, https://doi.org/10.1007/978-3-030-01437-7_16
- [39] V. Venkatesh, M. G. Morris, G. B. Davis, and F. D. Davis, “User acceptance of information technology: Toward a unified view,” *MIS Quarterly*, Vol. 27, No. 3, pp. 425-478, September 2003, <https://doi.org/10.2307/30036540>
- [40] V. Venkatesh and F. D. Davis, “A theoretical extension of the technology acceptance model: Four longitudinal field studies,” *Management Science*, Vol. 46, No. 1, pp. 186-204, February 2000, DOI:10.1287/mnsc.46.2.186.11926
- [41] J. Kang and S. Lee, “User-customized news service by use of social network analysis on artificial intelligence & bigdata,” *International Journal of Advanced Smart Convergence*, Vol. 10, No. 3, pp. 131-142, August 2021, <http://dx.doi.org/10.7236/IJASC.2021.10.3.131>

APPENDIX A

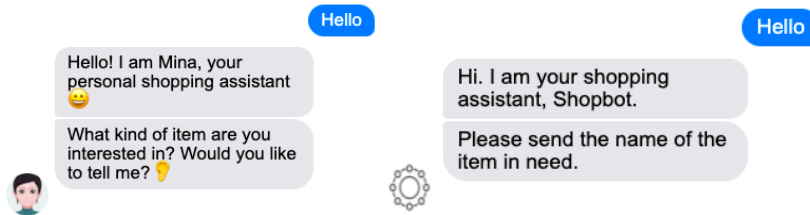


Figure A1. Sample chatbot conversation screenshot for high and low anthropomorphized conditions

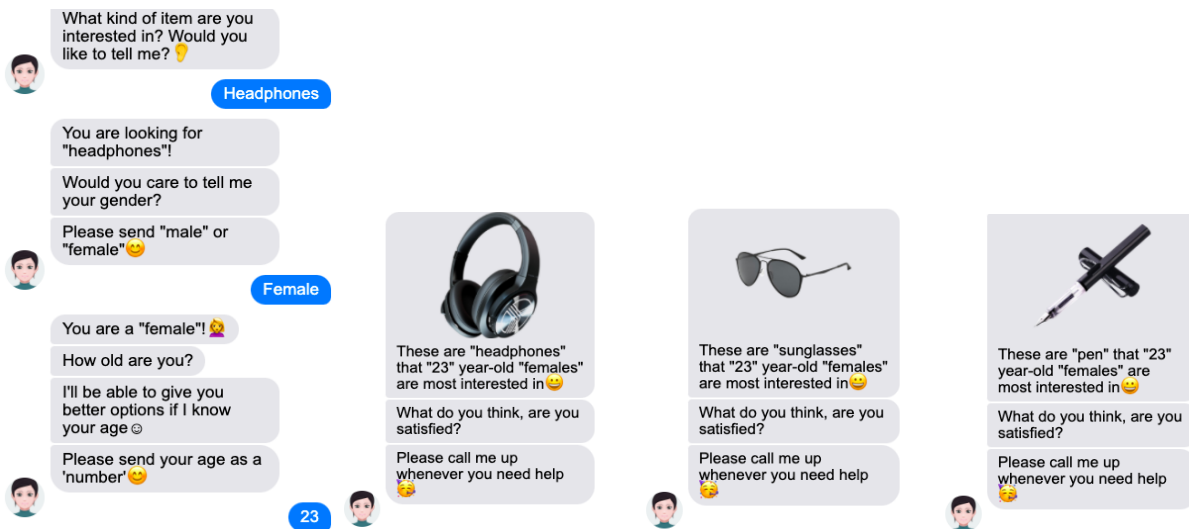


Figure A2. Chatbot conversation screenshot for high anthropomorphized conditions with high conversational relevance

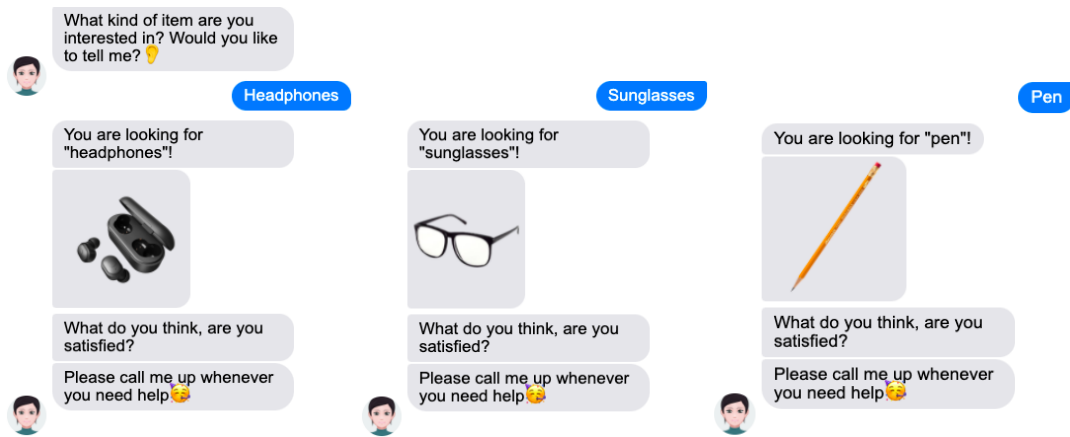


Figure A3. Chatbot conversation screenshot for high anthropomorphized conditions with low conversational relevance

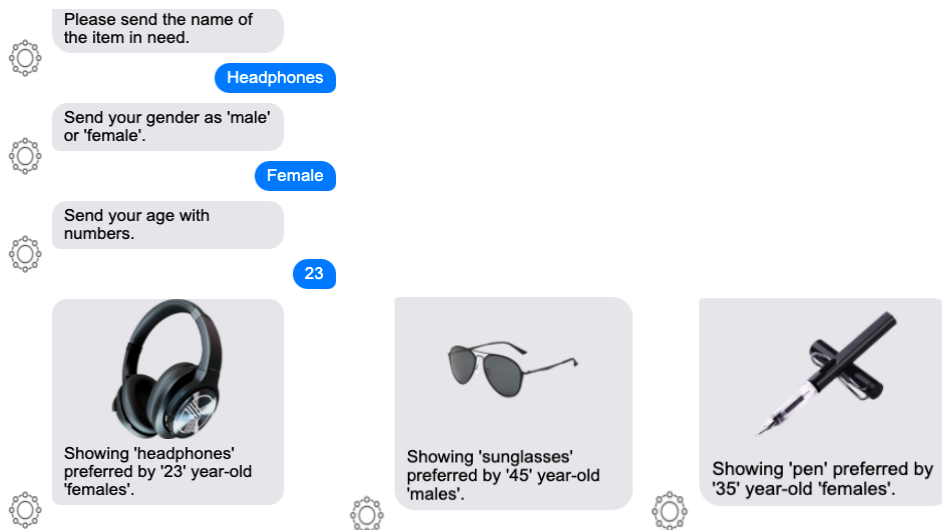


Figure A4. Chatbot conversation screenshot for low anthropomorphized conditions with high conversational relevance

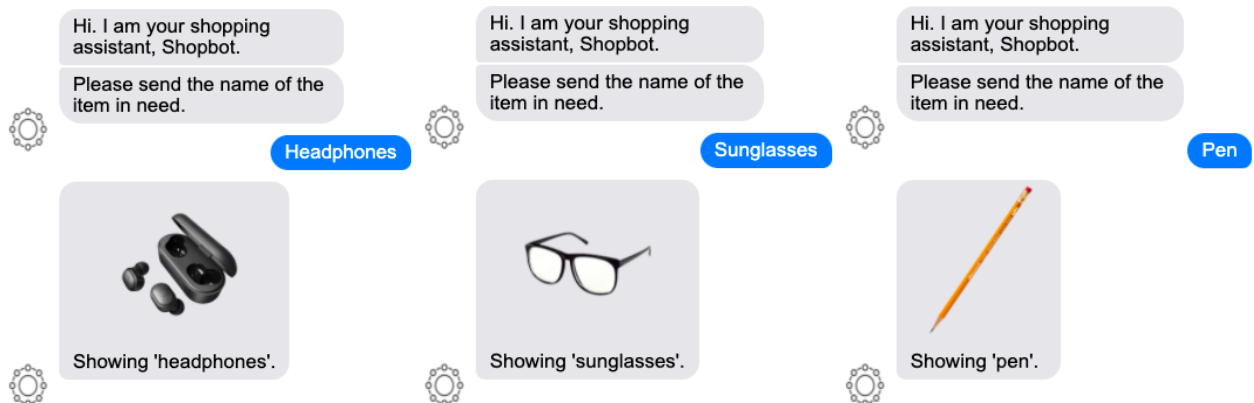


Figure A5. Chatbot conversation screenshot for low anthropomorphized conditions with low conversational relevance

APPENDIX B

Figure B1. Questionnaire items

Anthropomorphism	
Bartneck et al. (2009)	<ol style="list-style-type: none"> 1. Conversation with this chatbot feels natural. 2. This chatbot seems to have its own conscience. 3. This chatbot is humanlike. 4. This chatbot is lifelike.
Grimes et al. (2021)	<ol style="list-style-type: none"> 1. This chatbot is expected to give help to customers. 2. Conversation with this chatbot will be informative.
Manipulation check for conversational relevance	<ol style="list-style-type: none"> 1. This chatbot provided an accurate response for the consumer's needs.
Psychological distance (Li & Sung, 2021)	<ol style="list-style-type: none"> 1. This chatbot seems socially intimate with the consumer. 2. This chatbot seems psychologically intimate with the consumer. 3. This chatbot feels like a typical in-group member.
Usability (Bartneck et al., 2009)	<ol style="list-style-type: none"> 1. This chatbot is likeable. 2. This chatbot is skilled at what it is designed for. 3. This chatbot is smart. 4. This chatbot is responsible. 5. This chatbot has high levels of intelligence. 6. This chatbot is sensible.
Purchase intentions (Barber et al., 2012)	<ol style="list-style-type: none"> 1. If I were a consumer, I would consider purchasing the recommended item. 2. If I were a consumer, I am inclined to purchasing the recommended item. 3. If I were a consumer, I would plan on making a purchase based on the recommendation. 4. If I were a consumer, I have an intention to use the recommended item.