



# The effect of gender stereotypes on artificial intelligence recommendations

Jungyong Ahn<sup>a</sup>, Jungwon Kim<sup>b</sup>, Yongjun Sung<sup>b,\*</sup>

<sup>a</sup> School of Media & Communication, Korea University, Republic of Korea

<sup>b</sup> School of Psychology, Korea University, Republic of Korea

## ARTICLE INFO

### Keywords:

Artificial Intelligence (AI)  
AI agent  
Gender stereotypes  
AI recommendations

## ABSTRACT

This study explores the effects of gender stereotypes on evaluating artificial intelligence (AI) recommendations. We predict that gender stereotypes will affect human-AI interactions, resulting in somewhat different persuasive effects of AI recommendations for utilitarian vs. hedonic products. We found that participants in the male AI agent condition gave higher competence scores than in the female AI agent condition. Contrariwise, perceived warmth was higher in the female AI agent condition than in the male condition. More importantly, a significant interaction effect between AI gender and product type was found, suggesting that participants showed more positive attitudes toward the AI recommendations when the male AI recommended a utilitarian (vs. hedonic) product. Conversely, a hedonic product was evaluated more positively when advised by the female (vs. male) AI agent.

## 1. Introduction

We are entering the age of artificial intelligence (AI). AI represents human intelligence, including learning, reasoning, perceptual skills, and natural language understanding through computer programs (Brooks, 1991). Once just a popular trope in science fiction, AI technologies are now deeply embedded in and substantially changing every aspect of modern life. One of the fastest-growing areas of AI commercialization is e-commerce, particularly recommendation systems and AI agents that interact with consumers. AI agents, such as virtual assistants, AI speakers, and chatbots, deliver targeted advertising by linking with consumer databases such as Amazon and Google and responding immediately to users' inquiries about products and services. This enables companies to get a rapid turnaround without committing human resources. These advantages have made AI agents game-changers in online marketing that will evolve beyond even fiction writers' imaginations in the near future (Gentsch, 2018). For example, the number of consumers who shopped using Amazon's AI voice-activated in-home virtual assistant, Alexa, doubled from the 2017 to 2018 Christmas seasons. In China, the use of Alexa-type AI speakers and other AI-enabled voice assistants increased 42.7% in 2019 (Huang, 2020). Experts have predicted that consumer retail spending via chatbots will reach \$142 billion in 2021, up from \$2.8 billion in 2019 (Insider Intelligence, 2021).

People interact with diverse AI agents every single day. Consequently, the importance of interactions between humans and AI agents is

increasingly becoming pronounced in today's marketplace. Duffy (2003) predicted that developers would give robotic devices increasingly anthropomorphic features because many people distrust or dislike robots; this has certainly proved true for AI-driven e-commerce recommendation systems and virtual agents. *Anthropomorphism* is the human tendency to attribute intrinsic human properties and characteristics to animals and inanimate objects such as automobiles (Aggarwal & McGill, 2012; Epley et al., 2007). More than four decades ago, Weizenbaum (1976) said computers were just machines but humanized intelligence. Since many AI agents can convincingly simulate interpersonal interactions, it is not surprising that people interact with them as if they were humans. This shows a typical conceptual metaphor and the underlying process of anthropomorphism. Insights from research on interpersonal relationships (Cohrs et al., 2012; Fiske et al., 2007) and consumer-brand relationships (Fournier & Alvarez, 2012; Gao & Matila, 2014) suggest that human-AI agent interactions might be guided by the same norms governing interpersonal relationships.

This study suggests that gender stereotypes are a vital factor in how users associate AI agents with personality traits. People's perceptions and evaluations of others' personality traits affect their interpersonal communications. For example, perceiving a speaker as aggressive might make a listener respond defensively. Because people think of AI agents as having human characteristics and interact with them in ways that parallel social communication norms, they sometimes form trait inferences from their direct interactions with AI agents. These inferences

\* Corresponding author at: College of Law Annex 406, Korea University, Anam-dong 5 ga, Sungbuk-gu, Seoul, Republic of Korea.

E-mail address: [sungyj@korea.ac.kr](mailto:sungyj@korea.ac.kr) (Y. Sung).

<https://doi.org/10.1016/j.jbusres.2021.12.007>

Received 27 December 2020; Received in revised form 30 November 2021; Accepted 6 December 2021

Available online 15 December 2021

0148-2963/© 2021 Elsevier Inc. All rights reserved.

could color their perceptions of AI agents and their responses to the agents' recommendations. Research on stereotypes suggests that people are more likely to associate women with warmth and men with competence (Cuddy et al., 2008; Fiske et al., 2002). Studies have also shown that when individuals communicate with technological devices such as computers, machines, and AI entities, they use the same dialog styles and information-processing methods as they use in their human–human interpersonal communications (Kim et al., 2019; Nass et al., 1995; Nass & Moon, 2000).

Thus, we suggest that the gender stereotypes will also apply to human–AI interactions. We tested how the purported gender of AI agents was associated with the same two key personality traits used by Cuddy et al. (2008) and Fiske et al. (2002): competence and warmth. These are two traits people commonly consider when they evaluate others, and the traits are positively correlated with the level of trust in others (Fiske et al., 2002). We also tested the hypothesis that the consumers' reception to AI agents' product recommendations would be affected by their perceptions of the AI agents' "personalities" and their match with the product type being recommended (i.e., hedonic vs. utilitarian). To that end, we conducted two experiments. Experiment 1 examined the effect of gender stereotypes on consumers' reactions to (trust vs. no trust) a chatbots' product evaluations and recommendations. Experiment 2 used an AI speaker (e.g., Amazon's Alexa) to test the first experiment's findings and the mediation effect of warmth and competence.

## 2. Literature review

### 2.1. Interactions with AI

In e-commerce, AI can play various roles such as assistant, recommender, companion, and service agents. The increasingly common interactions between people and AI agents have attracted the interest of academic researchers. Many have investigated the effects of AI's features and user characteristics on interactions. For example, Edwards et al. (2019) found that students evaluated their AI instructor as having high credibility and social presence and found that older students rated older-sounding AI voices as more credible and motivating. Zhu and Deng (2021) examined the impact of social anxiety on adopting robotic training partners. They found that people with high social anxiety preferred robotic training partners over human training partners, finding the interactions less stressful. Furthermore, Fiore et al. (2013) examined the importance of robots expressing social cues in human–robot interaction. Social cues related to the robot's proxemic behavior positively affected people's perceptions of the robots' social presence and emotional state. However, cues related to the robot's gaze behavior had no impact.

Researchers are also increasingly investigating social cues and trait inferences in the context of AI as shopping agents. For example, Roy and Naidoo (2021) found that individuals with a present-time orientation preferred chatbot agents with a "warm" (personal) conversational style over those with a competency-focused (impersonal) style conversation. For those people, "warm" interactions led to positive product perceptions and greater purchase intentions. Chung et al. (2020) used a five-dimensional model measuring consumers' perceptions of such factors as interaction, entertainment, and problem-solving to investigate whether luxury brand retailers could still deliver personalized service with a chatbot. They found that the consumers considered the chatbots' e-service effective and "engaging" (p. 587).

The most commonly applied psychological theory used to explain human–AI (or robot) interactions is the *computers are social actors* (CASA) paradigm developed by Clifford Nass and others (Lee & Nass, 2005; Nass et al., 1994). The CASA paradigm suggests that in human–computer interactions, individuals unconsciously generate social responses, making those interactions fundamentally social. According to the CASA paradigm, individuals recognize computers as social actors

rather than machines (Lee & Nass, 2005; Nass et al., 1994; Nass et al., 1995; Reeves & Nass, 1996). Computers are inanimate objects; their perceived personalities are fictions created by computer programmers. Nonetheless, people infer electronic devices' personalities from the devices' speech patterns, vocabularies, and voice qualities, and they respond accordingly (Lee & Nass, 2005). The CASA paradigm perspective holds that individuals apply interpersonal communication norms to human–computer interactions. Byrne (1971) proposed what he called the *attraction paradigm*, which posits that people are attracted to attitudes similar to their own in interpersonal relationships. Similarly, Nass et al. (1995) found that people were more satisfied with and preferred to engage with computers they perceived as having personalities similar to theirs. The principle of reciprocity has also been applied to intimate self-disclosures with computer agents; individuals gave more personal information to computers when the computer disclosed its own information (e.g., technical specifications) than when it did not (Moon, 2000).

Researchers have also studied the stereotypes common in interpersonal relationships in the human–AI context. For example, Nass and Moon (2000) found that individuals relied on ethnic stereotypes in human–computer relationships. They asked participants to choose between two courses of action in a series of hypothetical choice-dilemma situations based on a computer agent's decision and arguments. One agent had an Asian avatar, the other a Caucasian avatar. The study found that the respondents' perceptions were affected by their ethnic stereotypes: the Asian participants found the Asian avatar more attractive, trustworthy, intelligent, and persuasive; the Caucasian participants preferred the Caucasian avatar. Similarly, Pratt et al. (2007) found that individuals were more persuaded by computer agents whose ethnicity was similar to theirs; individuals who received recommendations from computer agents they supposed to be of similar ethnicity were more likely to change their ranking based on computer agent's feedback (Pratt et al., 2007). In sum, while human–AI interactions are not identical to human–human interactions, studies of interpersonal relationships have proven useful for understanding the dynamics of the connections between humans and AI.

### 2.2. Gender stereotypes and personality traits

Stereotypes are generalized beliefs regarding specific groups (Heilman, 2012) and consciously or unconsciously intervene in our decision-making (Bargh et al., 1996; Dijksterhuis & Bargh, 2001; Hamilton et al., 1990). In particular, gender stereotypes are generalized attitudes toward men and women (Heilman, 2012), and individuals commonly apply them when forming impressions of others. For example, based on social role theory (Eagly, 2013), women are positioned as caregivers for children or the elderly more often than men. They are expected to show feminine behaviors such as nurturance or concern for personal relationships. In contrast, men are more likely to work outside the house, and accordingly, they are expected to behave in a masculine way, revealing assertiveness and leadership qualities (Eagly, 2013; Vogel et al., 2003). Similarly, Gupta et al. (2008) showed that, based on social stereotypes, men had a higher entrepreneurial intention, which is perceived to be a masculine trait, than women. Further, when one is adapting to a job, men were perceived to be better at managing stressful situations and adapting physically than women. Conversely, women were perceived to be better at learning and were expected to adapt interpersonally and culturally (DeArmond et al., 2006).

Prior research shows that gender stereotypes affect evaluating a target's competence and warmth (Broverman et al., 1972; Fiske et al., 2007; Williams & Best, 1990). When individuals judge others, they consider them based on two dimensions: competence and warmth (Fiske et al., 2007). Competence is a trait related to perceived ability, including intelligence, skill, and efficacy. Warmth is a trait related to perceived intent, including friendliness, helpfulness, and sincerity (Fiske et al., 2007). According to the results of previous studies, in general, individuals rate men as more competent than women and women as

warmer than men (Abele, 2003; Ashton-James et al., 2019; Eagly & Steffen, 1984; Heilman, 2012; Huddy & Terkildsen, 1993). Individuals predicted that male candidates would be more competent than female candidates, while female candidates would be warmer (Huddy & Terkildsen, 1993). Similarly, individuals viewed surgeons that are men as being more competent than surgeons that are women (Ashton-James et al., 2019).

As well as the evaluation of humans, previous studies found that individuals were affected by gender stereotypes during their interactions with computers. For example, according to Nass et al. (1997), participants showed gender-stereotypic responses toward computers with voices of different genders. More specifically, individuals perceived evaluations from a computer with a masculine voice as more valid than those from a computer with a feminine voice and considered dominant traits more desirable for men than women. Moreover, they expected computers that are men and women to know more about subjects typically regarded as masculine or feminine, respectively. Further, Tay et al. (2013) demonstrated that individuals evaluated a masculine robot as more suitable for a security task, which is stereotypically regarded as a masculine occupation, than a feminine robot. Further, Eysel and Hegel (2012) manipulated the gender of a robot by its appearance. They found that male robots were perceived as more agentic and better suited to masculine tasks requiring mathematical ability. In contrast, female robots were perceived as more communal and better suited to feminine tasks requiring verbal ability.

The results of previous studies suggest that such gender stereotypes apply when individuals interact with and evaluate AI agents. Thus, we predicted that individuals would use the same gender stereotypes when interacting with humans or AI agents and tested whether such interpersonal interaction norms would hold in human-AI interactions. Accordingly, the following hypothesis is advanced:

H1: Female AI will be perceived as warmer than male AI (a), whereas male AI will be perceived as more competent than female AI (b).

### 2.3. AI agent personalities and recommendations

The use of spokespersons, spoke characters, celebrity endorsers, and salespersons are very common in marketing, advertising, promotion, retail, and e-commerce. Personality traits associated with such people and characters are frequently incorporated into various business communications to persuade consumers. In particular, both competence and warmth are fundamental personality traits because they are the two key components and characteristics essential to trust (Sung & Kim, 2010).

Competence is related to the consumer's confidence that the product or service will provide quality performance in a reliable, responsible, and competent manner. It is the extent to which a brand is perceived as skillful and knowledgeable from experience or training in the product/service category. For example, Casaló et al. (2007) suggested that competence is a crucial element in e-commerce business. Consumers lack knowledge of the brands and sellers that operate in virtual environment. In addition, warmth positively affects trust and product/service evaluation (Kolbl et al., 2019). Consumers show higher loyalty to brands with higher warmth, and warmth contributes to forming long-term relationships between consumers and brands (Fournier & Alvarez, 2012; Malone & Fiske, 2013). Consumers' beliefs about people being warm, sincere, and honest are critical factors that increase brand trust (e.g., Chaudhuri & Holbrook, 2001). Thus, if gender stereotypes affect the evaluation of AI agents in terms of their personality traits (as proposed in H1), consumers will perceive male AI more competently and female AI more warmly, and the personalities of AI agents will affect product evaluation.

In the present study, we proposed that the two personality traits' effects on evaluating AI-recommended products will vary by product type (utilitarian vs. hedonic; Bennett & Hill, 2012; Chattalas & Takada, 2013). Utilitarian products are products for practical or functional

purposes, while hedonic products relate to emotional experiences such as pleasure and sadness (Holbrook & Hirschman, 1982). When consumers judge whether a utilitarian product is practical, the product's competence (e.g., product functions and performance) has a crucial effect on their judgment (Ahn et al., 2020; Dhar & Wertenbroch, 2000). As the spokesperson's competence can represent the competence of the product (Parulekar & Raheja, 2006), the spokesperson's competence is an essential factor in purchasing utilitarian products, where competence is a necessary virtue of product selection (Bennett & Hill, 2012). In contrast, for hedonic products, consumers' emotional experience or subjective feeling is significant in helping them decide whether to buy a product (Ahn et al., 2020; Dhar & Wertenbroch, 2000). Consumers' affective responses affect their evaluations of hedonic products (Yeung & Wyer, 2004). The warmer consumers perceive the spokesperson to be, the more consumers will respond positively to the spokesperson, which also affects the sale of the products (Bennett & Hill, 2012; Chattalas & Takada, 2013). In contrast to utilitarian products that require rational judgment, emotional judgment has an essential effect on sales in hedonic products. Hence, the spokesperson's warmth, which can positively affect consumers' emotional decision, has a more critical impact than competence on the successful sale of hedonic products.

Thus, if the gender of AI affects how consumers perceive the AI agent's personalities (competence and warmth), such interactions between AI agents' personalities and product type (utilitarian and hedonic) will affect the evaluation of an AI's recommendations. The following hypotheses are proposed:

H2: For utilitarian products, male AI will be more persuasive than female AI (a), whereas, for hedonic products, female AI will be more persuasive than male AI (b).

H3: The personality traits (competence and warmth) of AI will mediate the interaction between the gender of AI and the product type, which will affect the evaluation of AI's recommendations.

Fig. 1 shows a conceptual model of our proposed hypotheses. To test our hypotheses, we conducted two experimental studies. In Experiment 1, we created three different gender types of AI chatbots to test the effects of gender stereotypes on the personality traits (i.e., competence and warmth) of AI (H1-a and H1-b). Further, we tested the hypotheses (H2-a and H2-b) that the AI chatbot's expertise in product types (utilitarian and hedonic) perceived by individuals varies according to the gender of the AI chatbot. In Experiment 2, we created a real communication environment between an AI speaker and a user and replicated the findings of Experiment 1. In addition, we demonstrated the effect of the personalities of AI as the mediator of the interaction effect between AI gender and product type (H3).

## 3. Experiment 1

### 3.1. Materials and methods

Experiment 1 was designed to prove the effect of gender stereotypes on AI personalities (H1) and the interaction effect between AI gender and product type on evaluating AI's recommendations (H2). We used a 3 (AI gender: male vs. female vs. neutral)  $\times$  2 (product type: utilitarian vs. hedonic) between-subjects design. In Experiment 1, we created a set of AI chatbots according to the gender condition (see appendix 1). To select both utilitarian and hedonic products, a pre-test was conducted. We showed the participants a total of six products and then asked the buying purpose of these items (1 = functional, practical, unenjoyable, not fun; 7 = not functional, impractical, enjoyable, fun; Crowley et al., 1992). As a result of the pre-test with 40 participants, the wireless headphone was the most suitable for the utilitarian product. The drawing tool was the most ideal for the hedonic product (wireless headphone:  $M = 2.82$ , drawing tool:  $M = 5.82$ ,  $p < .001$ ).

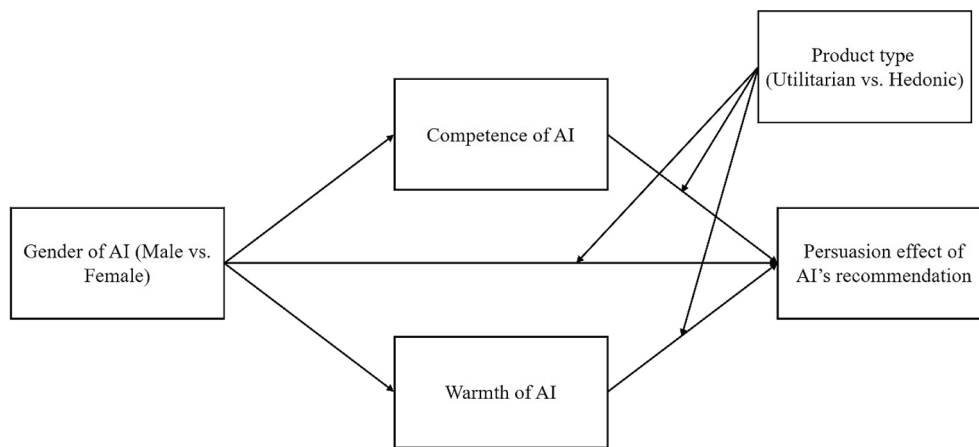


Fig. 1. Conceptual Model of the Proposed Hypotheses.

### 3.2. Participants and procedure

Experiment 1 was conducted through an online experiment. Participant were recruited through the online community of university located in Seoul, South Korea. It was informed that participants would receive about \$5 worth of coupon if they participate in the experiment and students who wish to participate in the experiment were provided online link. A total of 180 undergraduates (age  $M = 23.91$ ) participated and the male-to-female ratio was 50:50 under all conditions to control for the effect of participants' gender. Experiment 1 was divided into two sessions. In the first session, participants accessed the chat system through the researcher's link and asked the chatbot to introduce itself. Then the chatbot revealed its name, occupation, and gender (male, female or neutral). Male and female chatbots said their gender is either male or female, while the neutral chatbot said, "I have no gender because I'm an AI." After that, the participants asked the chatbot about the current domestic coronavirus disease (COVID-19) situation. The chatbot informs about the virus' situation in real-time.

In the second session, the participants asked the chatbot to show the latest news related to COVID-19, and the chatbot showed the article and insisted that wireless headphones (or drawing tools) are attracting attention in the COVID-19 era depending on product type conditions. In the utilitarian product condition, the article said that wireless headphones' sales had increased rapidly due to the recent increase in telecommuting working out of the home. For hedonic products, the article was provided to the participants stating that drawing was gaining interest as an indoor hobby due to restrictions on going out. After reading the article, the participants asked the chatbot how to purchase the product wisely. Then the chatbot gave tips about priorities to consider when purchasing the product. After the conversation, the participants were asked to answer a series of questions to evaluate the chatbot.

### 3.3. Measures

#### 3.3.1. Manipulation checks

In the last section of the experiment, the participants were asked to identify the gender of the chatbot with which they interacted. A total of 17 participants could not correctly answer the question on the chatbot's gender. Therefore, only 180 out of 197, excluding 17, were used for the data analysis. They also answered whether the product the chatbot mentioned was utilitarian or hedonic (1 = *functional, practical, unenjoyable, not fun*; 7 = *not functional, impractical, enjoyable, fun*;  $\alpha = 0.71$ ). As a result, the score of drawing tools ( $M = 5.80$ ) was significantly higher than the score of wireless headphones ( $M = 2.91$ ;  $p < .001$ ).

#### 3.3.2. AI personalities

The instrument to measure the personality of the AI scored six items

on a 7-point Likert scale (1 = *strongly disagree*, 7 = *strongly agree*), comprising three items that measured competence (i.e., *competent, intellectual, skillful*;  $\alpha = 0.85$ ) and three items that measured warmth (i.e., *warm, friendly, thoughtful*;  $\alpha = 0.77$ ; Fiske et al., 2007).

#### 3.3.3. Dependent variable

The AI's expertise in the product (wireless headphones or drawing tool) was measured with three items ("What do you think is the chatbot's expertise in the product": 1 = *very poor*, 7 = *very good*; "How much can you trust the chatbots' buying tips?": 1 = *never trust*, 7 = *highly trust*; "If you're buying the product, are you willing to use the buying tips from the chatbot?": 1 = *never use*, 7 = *very positively*;  $\alpha = 0.90$ ).

#### 3.3.4. Control variable

Participants' involvement of wireless headphones or drawing tool was used as a control variable (knowledge of the product: 1 = *very low*, 7 = *expert level*; interest of the product: 1 = *not interest at all*, 7 = *very interest*;  $\alpha = 0.81$ ).

### 3.4. Results

To test H1-a and H1-b, a one-way ANOVA was conducted. The results showed that warmth of AI differed depending on the genders of AI ( $F(2, 177) = 13.54$ ,  $p < .001$ ,  $\eta_p^2 = 0.13$ ). As a result of post-hoc analysis (LSD), the warmth of female AI ( $M = 4.71$ ) was significantly higher than that of male ( $M = 3.79$ ) and neutral AI ( $M = 4.07$ ;  $p < .001$ ), and the difference between men and neutral was not significant ( $p = .13$ ), supporting H1-a. However, regarding competence (H1-b), AI gender differences were not founded across the three conditions (male:  $M = 4.64$ ; female:  $M = 4.43$ ; neutral:  $M = 4.44$ ;  $ps < 0.05$ ).

A  $3 \times 2$  ANCOVA was conducted to test the interaction between AI gender and product type on the AI's expertise. The main effects of AI gender ( $p = .28$ ) and product type ( $p = .44$ ) were found to be not significant. However, the interaction effect of AI gender  $\times$  product type was statistically significant ( $F(2, 174) = 11.98$ ,  $p < .001$ ,  $\eta_p^2 = 0.12$ ). Subsequent contrast analyses showed that for the utilitarian product, participants evaluated the male AI ( $M = 5.08$ ) more highly than that of female ( $M = 4.40$ ) and neutral conditions ( $M = 4.54$ ) ( $ps < 0.001$ ). The difference between female and neutral AI was not significant though ( $p = .64$ ).

In the case of the hedonic product, however, female AI expertise ( $M = 5.17$ ) was the highest, followed by neutral AI ( $M = 4.46$ ) and male AI ( $M = 4.04$ ) ( $ps < .001$ ). There was no statistically significant difference between male and neutral AI ( $p = .07$ ). Finally, additional analyses showed that for male AI, expertise in the utilitarian product was higher than that of the hedonic product ( $p < .01$ ). In contrast, for female AI, expertise in the hedonic product was more highly evaluated than the



utilitarian product ( $p < .01$ ). In the case of neutral AI, there was no difference in the expertise of AI between utilitarian and hedonic products ( $p = .83$ ). These results support H2a and H2b (see Fig. 2).

### 3.5. Discussion in brief

Experiment 1 provided empirical evidence for our hypotheses on the effect of gender stereotypes on AI evaluations. The results showed that the participants' perceptions of the AI agent's gender significantly affected the participants' evaluations of the AI agents. In line with our predictions, the participants perceived that the "female" AI agent was warmer than the male agent or the neutral agent (H1-a). However, contrary to our predictions, we did not find a significant difference in the participants' perceptions of the AI agents' competence for the three genders (H1-b). Contrast analysis results showed that the participants in the male AI  $\times$  utilitarian product condition ( $M = 5.07$ ) perceived the AI agent as more competent than participants in the male AI  $\times$  hedonic product condition ( $M = 4.22$ ;  $p < .001$ ). For the female AI agent, the participants in the hedonic product condition ( $M = 4.77$ ) perceived the competence of the AI agent as higher than those in the utilitarian product condition ( $M = 4.09$ ;  $p < .05$ ). This suggests that the matching effect between the AI agents' perceived gender and the product type significantly influenced the participants' perception of the AI agents' personalities. In Experiment 1, the interaction effect between AI gender and product type was verified (H2a, b). For utilitarian products, the participants evaluated the expertise of the male AI agent higher than the female AI agent; for the hedonic product, the results were the opposite. In the neutral AI conditions, however, we found no difference in the participants' evaluations of the AI agents according to product type. Thus, our predictions that AI gender influences AI evaluations was firmly supported.

In Experiment 2, we designed a separate personality evaluation session to independently confirm the effect of AI agents' purported gender on the participants' perceptions of the AI agents' personalities. We also tested the proposed mechanism that perceptions of AI agents' personalities caused by gender stereotypes mediated the interaction effect between the AI agents' perceived gender and the product type in the participants' evaluations of the agents' recommendations.

## 4. Experiment 2

### 4.1. Materials and methods

Experiment 2 aimed to replicate the result of Experiment 1 (H1 and 2), and also test our H3. We used a 2 (AI gender: male vs. female)  $\times$  2

(product type: utilitarian vs. hedonic) between-subjects design. In Experiment 1, an AI chatbot was employed, whereas, in Experiment 2, an AI speaker was used. To manipulate AI gender, we recorded real human voices. To control for the effect of the attitude toward the voice itself rather than its gender, we recorded the voices of four male and four female speakers in advance and examined perspectives toward the voices. A total of 80 participants were recruited for the pre-test, and 10 participants evaluated each voice. The participants listened to the agent for 15 s and considered the gender accuracy of the voice and attitude toward the voice (1 = *bad/negative*; 7 = *good/positive*). As a result of the pre-test, the pair of male and female voices (that were identified with 100% accuracy) with the most insignificant differences in attitude score (male voice:  $M = 4.70$ , female voice:  $M = 4.65$ ,  $p < .98$ ) was selected. Further, we pre-tested with an independent group of 40 participants to choose utilitarian and hedonic products. Six items (three utilitarian products: ergonomic mouse, thermos, portable charger; three hedonic products: scented candle, chocolate box, unisex bracelet) that cost approximately \$20 were rated in terms of the buying purpose of these items (1 = *functional, practical, unenjoyable, not fun*; 7 = *not functional, impractical, enjoyable, fun*; Crowley et al., 1992) and attitudes toward them (1 = *bad, negative*; 7 = *good, positive*). The results showed an ergonomic mouse was the most suitable utilitarian item and a scented candle the most suitable hedonic item (ergonomic mouse:  $M = 2.45$ , scented candle:  $M = 5.94$ ,  $p < .001$ ).

### 4.2. Participants and procedure

A total of 120 undergraduates (age  $M = 23.4$ ) participated in the laboratory experiment. The male-to-female ratio was 50:50 under all conditions to control for the effect of participants' gender. We used the Wizard of Oz method (Dahlbäck et al., 1993) to make participants believe they were interacting with a real AI speaker instead of a manipulated one as a control. Once each participant entered the laboratory, they were informed that the AI agent was a newly developed AI speaker. The experiment was divided into two sessions. In the first session, the participants were asked to interact with the assigned AI speaker (either male or female) for approximately 3 min. The participants provided their necessary demographic information, such as name, age, gender, and academic status, to the AI speaker. The AI speaker briefly introduced himself or herself to the participants. At the end of this conversation, the AI speaker said that the participants' voices and demographic information were fully recognized. It would serve as the personal AI assistant for the participants. After the interaction with the AI agent, the participants were asked to evaluate its personality.

In the second session, participants in each condition were instructed

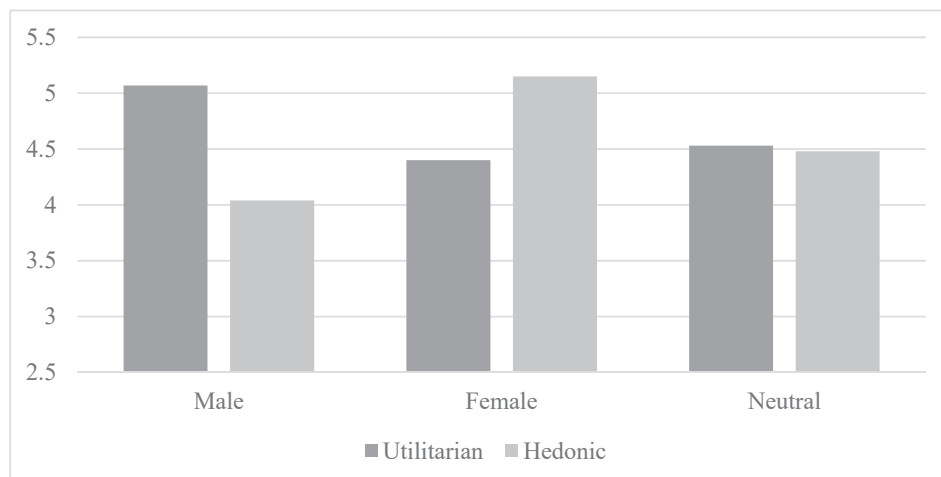


Fig. 2. AI Gender  $\times$  product type on AI's expertise (Experiment 1).

to ask the AI speaker to recommend a popular item for college students. For the utilitarian product condition, the AI speaker said, “Ergonomic mice are among the most popular items for college students.” Conversely, under the hedonic product condition, participants heard, “Scented candles are among the most popular items for college students.” Subsequently, the participants for both products were recommended by the AI agent. During the process, the AI agent sounded several repetitive alarms. After saying, “Matching the results from the big data, I found the perfect item for you,” it recommended the item (mouse or candle) with a 15-second commentary. After the interactions, the participants were asked to answer a series of questions to evaluate the AI’s recommendations.

### 4.3. Measures

#### 4.3.1. Manipulation checks

In the last section of the experiment, the participants were asked to identify the gender of the AI speaker with which they interacted. All participants identified the gender of the AI correctly. Further, they answered whether the product the AI recommended was utilitarian or hedonic. Questions were then asked regarding four characteristics, scored on a 7-point Likert-type scale (1 = *functional, practical, unenjoyable, not fun*; 7 = *not functional, impractical, enjoyable, fun*;  $\alpha = 0.75$ ; Crowley et al., 1992). The participants rated the scented candle ( $M = 5.72$ ) higher than the ergonomic mouse ( $M = 2.38$ ;  $p < .001$ ). Hence, all independent variables were successfully manipulated.

#### 4.3.2. AI personalities

The instrument to measure the AI speaker’s personality scored six items on a 7-point Likert scale (1 = *strongly disagree*, 7 = *strongly agree*), comprising three items that measured competence (i.e., *competent, intellectual, skillful*;  $\alpha = 0.80$ ) and three items that measured warmth (i.e., *warm, friendly, thoughtful*;  $\alpha = 0.79$ ; Fiske et al., 2007).

#### 4.3.3. Dependent variables

The evaluations of the AI’s recommendations were measured along the dimensions of product attitude and purchase intention (product attitude: 1 = *negative, bad, unfavorable*; 7 = *positive, good, favorable*;  $\alpha = 0.93$ ; purchase intention: 1 = *not buy, unnecessary*; 7 = *buy, necessary*;  $\alpha = 0.83$ ).

#### 4.3.4. Control variable

Participants’ product (ergonomic mouse or scent candle) involvement was used as a control variable (knowledge of the product: 1 = very low, 7 = expert level; interest of the product: 1 = not interest at all, 7 = very interest;  $\alpha = 0.63$ ).

### 4.4. Results

Frist, we used a  $2 \times 2$  ANCOVA to replicate H2 (the interaction effect between the gender of AI and the product type on the dependent variables). There was no significant main effect for the gender of AI (product attitudes,  $p = .25$ ; purchase intention,  $p = .94$ ) or the product type (product attitudes,  $p = .64$ ; purchase intention,  $p = .92$ ). Control variable (product involvement) also has not statistically significant impact on dependent variables (product attitudes,  $p = .73$ ; purchase intention,  $p = .81$ ). However, the effect of the AI gender  $\times$  product type interaction was significant for both product attitudes ( $F(1, 116) = 6.25, p < .01, \eta_p^2 = 0.05$ ) and purchase intention ( $F(1, 116), p < .05, \eta_p^2 = 0.04$ ). Subsequent contrast analyses showed that participants favored the utilitarian product when the male AI (product attitudes:  $M = 4.54$ ; purchase intention:  $M = 3.97$ ) recommended it than when the female AI recommended it (product attitudes,  $M = 3.91, p < .01$ ; purchase intention:  $M = 3.51, p < .05$ ). In contrast, when the recommended product was hedonic, the female AI induced a more positive score (product attitudes:  $M = 4.69$ ; purchase intention:  $M = 4.00$ ) than the male AI (product

attitudes:  $M = 4.26, p < .01$ ; purchase intention:  $M = 3.52, p = .05$ ) (see Fig. 3).

To confirm our H1 and H3, we used an SPSS PROCESS version 3.5 created by Hayes to estimate the conditional indirect effect (Hayes & Scharkow, 2013). We adopted the 15th model from Hayes’s templates (2017) for statistical analysis. The independent variable was the gender of the AI (male vs. female), with the male and female AI coded as  $-1$  and  $1$ , respectively. The mediator was the AI’s two personalities: competence and warmth. The moderator was the product type (utilitarian vs. hedonic), with utilitarian and hedonic products coded as  $-1$  and  $1$ , respectively. Finally, the dependent variables were product attitudes and purchase intention. Statistical tests included a series of bootstrap analyses with 5000 samples and 95% bias-corrected confidence intervals (Preacher et al., 2007). The analysis showed that the gender of AI affects the evaluation of the AI’s competence ( $B = -0.61, SE = 0.16$ ) ( $p < .001$ ) and warmth ( $B = 1.32, SE = 0.25; p < .001$ ). An additional  $t$ -test suggested the same patterns. Participants perceived the male AI ( $M = 5.26$ ) as more competent than the female AI ( $M = 4.62; p < .001$ ). Conversely, they gave higher scores on warmth to the female AI ( $M = 5.03$ ) than to the male AI ( $M = 3.73; p < .001$ ). In addition, we found that the conditional indirect effect was statistically significant. First, the direct effect of the gender of the AI on product attitude was not significant ( $p = .24$ ). Additionally, the interaction effect between the gender of the AI and the product type was not significant ( $p = .12$ ). The interaction effect between the personalities (competence and warmth) of the AI and the product type on product attitude, however, was statistically significant (competence  $\times$  product type:  $B = 0.72, SE = 0.19, p < .01$ ; warmth  $\times$  product type:  $B = -0.79, SE = 0.12, p < .01$ ). More specifically, the conditional indirect effect of AI competence was significant for the utilitarian product ( $B = 0.77, SE = 0.14, p < .01$ ; 95% CI = 0.50 to 1.05), but not for the hedonic product ( $p = .69$ ). In contrast, AI warmth positively affected product attitudes only when the AI recommended the hedonic product ( $B = 0.67, SE = 0.08, p < .001$ ; 95% CI = 0.52 to 0.83), not the utilitarian product ( $p = .17$ ; 95% CI =  $-0.30$  to  $0.05$ ). For another dependent variable, purchase intention, the mediation analysis showed the same pattern for product attitudes (see Table 1). Overall, these results support H1 and H3.

## 5. General discussion

Given the proliferation of AI agents, it is vital to understand the factors affecting AI agents’ persuasiveness. Our study focused on AI agents providing e-commerce recommendations. We proposed that the purported gender of an AI agent would affect participants’ perceptions of its personality such that the AI agents’ persuasiveness would differ according to that gender. We demonstrated that gender stereotypes (as described in the interpersonal interaction literature) applied to the AI–human interactions and affected the participants’ evaluations of the AI agents’ recommendations. Specifically, they perceived the “male” AI agent to be more competent than the “female” one and the “female” AI agent to be warmer than the “male” one.

Furthermore, the persuasiveness of the AI agents’ recommendations varied according to their perceived personalities. For the utilitarian products, the participants’ trusted the “male” AI agents’ recommendations more than the “female” AI agents’ recommendations. The opposite was true for the hedonic products. Thus, we concluded that the perceived personalities of the AI agents mediated the effect of the AI agents’ purported gender and the product type on the participants’ responses to the AI agents’ recommendations.

Our research has both theoretical and practical implications. First, as AI agents have become commonplace in our daily lives, researchers have increasingly investigated the factors influencing AI–human interactions, such as conversation style (Roy & Naidoo, 2021), level of anthropomorphism (Pizzi et al., 2021), and social role (Rhee & Choi, 2020). Previous research has examined the effects of AI agents’ purported gender and personality on human–AI interactions (Kim et al., 2019; Nass

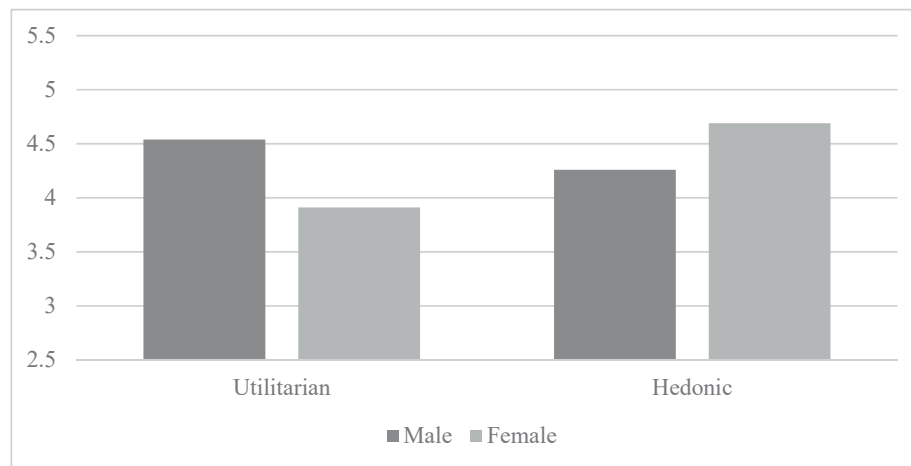


Fig. 3. AI Gender × product type on product attitude (Experiment 2).

Table 1

Conditional indirect effect of gender of AI × product type on dependent variables (Experiment 2).

| First Step: AI's personalities   |                   |      |        |                |  |                    |      |                |                |
|----------------------------------|-------------------|------|--------|----------------|--|--------------------|------|----------------|----------------|
|                                  | Competence        |      |        |                |  |                    |      |                |                |
|                                  | B                 | SE   | p      | CI             |  |                    |      |                |                |
| Gender of AI (X)                 | −0.61             | 0.16 | <0.001 |                |  |                    |      | −0.92 to −0.29 |                |
| Product involvement (CO)         | −0.01             | 0.08 | 0.939  |                |  |                    |      | −0.16 to 0.15  |                |
|                                  | Warmth            |      |        |                |  |                    |      |                |                |
|                                  | B                 | SE   | p      | CI             |  |                    |      |                |                |
| X                                | 1.32              | 0.25 | <0.001 |                |  |                    |      | 0.83 to 1.81   |                |
| CO                               | −0.07             | 0.12 | 0.562  |                |  |                    |      | −0.31 to 0.17  |                |
| Second Step: Dependent Variables |                   |      |        |                |  |                    |      |                |                |
|                                  | Product Attitudes |      |        |                |  | Purchase Intention |      |                |                |
|                                  | B                 | SE   | p      | CI             |  | B                  | SE   | p              | CI             |
| X                                | −0.66             | 0.56 | 0.242  | −1.78 to 0.45  |  | −0.89              | 0.60 | 0.140          | −0.92 to −0.29 |
| Product type (MO)                | −0.77             | 1.26 | 0.540  | −3.26 to 1.72  |  | −0.10              | 1.33 | 0.940          | −2.75 to 2.54  |
| Competence (ME1)                 | −0.67             | 0.29 | <0.05  | −0.12 to −0.10 |  | −0.31              | 0.31 | 0.316          | −0.92 to 0.30  |
| Warmth (ME2)                     | 1.47              | 0.18 | <0.001 | 1.11 to 1.83   |  | 1.38               | 0.19 | <0.001         | 1.00 to 1.76   |
| X × MO                           | 0.56              | 0.35 | 0.117  | −0.14 to 1.26  |  | 0.58               | 0.38 | 0.128          | −0.17 to 1.32  |
| X × ME1                          | 0.72              | 0.19 | <0.001 | 0.35 to 1.10   |  | 0.51               | 0.20 | <0.01          | 0.11 to 0.91   |
| X × ME2                          | −0.79             | 0.12 | <0.001 | −1.03 to −0.56 |  | −0.75              | 0.13 | <0.001         | −1.00 to −0.50 |
| CO                               | −0.07             | 0.08 | 0.351  | −0.22 to 0.08  |  | −0.04              | 0.08 | 0.654          | −0.13 to 0.20  |

et al., 1997; Tay et al., 2013). However, they have not empirically shown how perceptions of AI agents' gender affect consumer behavior. Our research provides empirical evidence that because they rely on gender stereotypes, consumers' trust levels differ depending on the perceived gender of AI agents providing e-commerce product recommendations. Furthermore, we examined the interactive effect of the AI agents' perceived gender and the product type and identified a psychological mechanism: the participants' perceptions of the AI agents' personalities (based on gender stereotypes) mediated the interaction effects between the two variables. We showed that the social psychology theory on gender stereotypes (Fiske et al., 2007) could be applied to new domains by revealing that gender stereotypes affect AI–human interactions and consumers' trust in AI agents' recommendations. This finding suggests that AI agents' "gender" and "personalities," influenced by people's gender stereotypes, affect consumers' receptions of AI recommendation systems in the shopping context.

Second, we established our studies' validity through the research design. Our experiments used the latest, most popular AI technology currently available—chatbots (Experiment 1) and virtual assistant devices (Experiment 2) that consumers interact with daily in a various

commercial and personal settings. We used the most readily available and convincingly simulated interpersonal interaction resources in both offline and online environment environments. Therefore, we believe our study's findings to have high external validity.

Third, our study provides some important managerial guidelines for commercial enterprises considering the adoption of AI agents. Our findings suggest that they should consider the purported gender of their AI agents to ensure that they deliver the most persuasive recommendations for specific products. For example, our findings suggested that consumers purchasing travel products would be more satisfied with recommendations from AI agents they perceived as female (vs. male) because of the hedonic nature of travel products and services and consumers' stereotypical association between women and warmth (Chatalas & Takada, 2013). In contrast, for electronic shops, whose products are utilitarian, consumers would be more satisfied with recommendations from AI agents they perceived as male, given their stereotypical association between men and competence. Furthermore, the product type could be manipulated differently for the same products according to sales strategies. For example, consider recommendations for jeans: if the comfort is emphasized, consumers perceive the jeans as utilitarian,

but if the style is emphasized, they perceived them as hedonic (Adaval, 2001). Therefore, if marketers can manipulate the product type according to the sales strategy and set the gender of the AI agent accordingly, they should see increased sales.

Finally, for commercial applications, different AI agents' avatars can be designed with various combinations of attributes to achieve different goals. Recently, Amazon announced that the voice of popular actor Samuel L. Jackson would be available for its Alexa device allows users to instruct Alexa to tell jokes, sing happy birthday, announce the weather, and more with his voice—including optional profanity. (Presumably, other celebrity voice options will be available that will offer other options.) AI agents' "personality" attributes such as gender, ethnicity, and age (and, in the case of the actor, associations with his movie characters' lines) can be reflected in the voice. Therefore, the voice is likely to trigger various stereotypes accordingly (Zuckerman & Driver, 1989). For example, individuals expect the elderly to be empathetic and friendly (Kornadt, 2016), be wise, and value social harmony; thus, a counselling AI agent might be given an older person's voice to provide a comforting and trustworthy impression. We recommend that developers consider their AI agents' voice characteristics and select those characteristics most appropriate proper for their target users.

This study had some limitations. First, this study relied on a limited number of product types and personality traits. Further research (with more personality attributes and product types) is needed to explore the generalizability of these results. Additionally, it is not only products that can be divided into hedonic and utilitarian types. This can also be done with places (Park, 2004), services (Pizzi et al., 2021), and shopping experiences (Scarpi, 2021). Thus, future studies should extend this study to other contexts and subjects.

Second, future research is needed to replicate our study with other types of stereotypes such as nationality, race, age, occupation (Fiske et al., 2007), religion, sexual identity, social status, and culture. There is

a need for a cross-cultural version of our study that examines the effect of gender and other stereotypes in human–AI interactions. For example, in Nordic countries with a high gender equality index (World Economic Forum, 2019), gender stereotypes might be relatively weaker than in other countries, leading to different results. Additional cross-cultural research will provide insights into the separate roles that cultural origin and gender stereotypes play in the perception and evaluation of AI agents' trustworthiness. Additional studies could enhance our understanding of the consequence of different approaches to AI technology development and design across cultures.

#### CRediT authorship contribution statement

**Jungyong Ahn:** Conceptualization, Data curation, Experiment design, Formal analysis, Funding acquisition, Investigation, Methodology, Writing – original draft, Writing – review & editing. **Jungwon Kim:** Conceptualization, Experiment design, Formal analysis, Investigation, Visualization, Validation, Writing – original draft, Writing – review & editing. **Yongjun Sung:** Conceptualization, Investigation, Supervision, Validation, Visualization, Writing – review & editing.

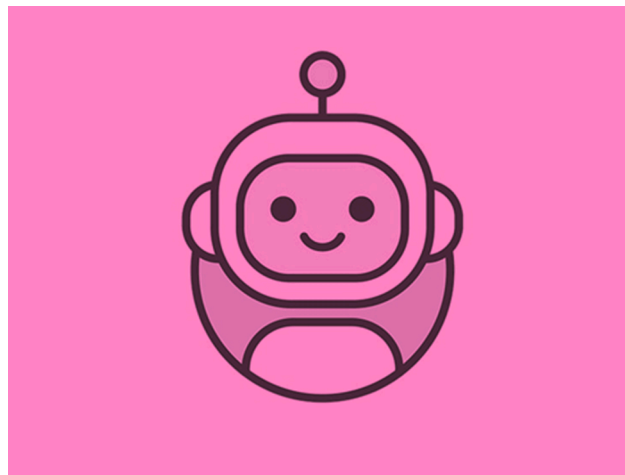
#### Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

#### Acknowledgments

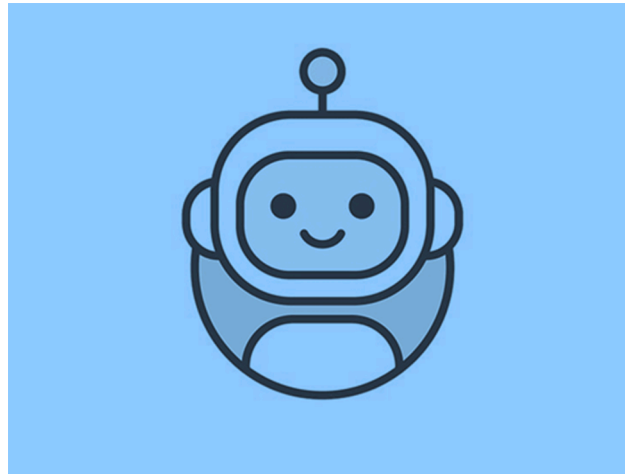
This work was supported by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NRF-2021S1A5B5A16075775).

## Appendix 1

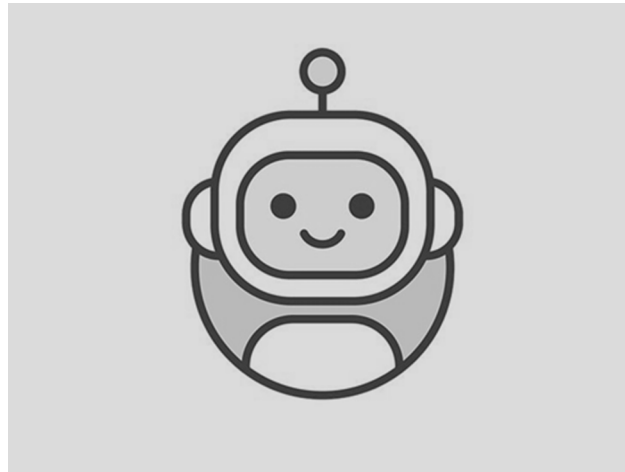


AI chatbot in the female condition





AI chatbot in the male condition



AI chatbot in the neutral condition

## References

- Abele, A. E. (2003). The dynamics of masculine-agentic and feminine-communal traits: Findings from a prospective study. *Journal of Personality and Social Psychology*, 85(4), 768–776.
- Adaval, R. (2001). Sometimes it just feels right: The differential weighting of affect-consistent and affect-inconsistent product information. *Journal of Consumer Research*, 28(1), 1–17.
- Aggarwal, P., & McGill, A. L. (2012). When brands seem human, do humans act like brands? Automatic behavioral priming effects of brand anthropomorphism. *Journal of Consumer Research*, 39(2), 307–323.
- Ahn, J., Kim, A., & Sung, Y. (2020). The effects of sensory fit on consumer evaluations of co-branding. *International Journal of Advertising*, 39(4), 486–503.
- Ashton-James, C. E., Tybur, J. M., Grieser, V., Costa, D., & Lindeman, B. (2019). Stereotypes about surgeon warmth and competence: The role of surgeon gender. *PloS One*, 14(2), e0211890. <https://doi.org/10.1371/journal.pone.0211890>
- Bargh, J. A., Chen, M., & Burrows, L. (1996). Automaticity of social behavior: Direct effects of trait construct and stereotype activation on action. *Journal of Personality and Social Psychology*, 71(2), 230–244.
- Bennett, A. M., & Hill, R. P. (2012). The universality of warmth and competence: A response to brands as intentional agents. *Journal of Consumer Psychology*, 22(2), 199–204.
- Brooks, R. A. (1991). Intelligence without representation. *Artificial Intelligence*, 47(1–3), 139–159.
- Broverman, I. K., Vogel, S. R., Broverman, D. M., Clarkson, F. E., & Rosenkrantz, P. S. (1972). Sex-role stereotypes: A current appraisal 1. *Journal of Social Issues*, 28(2), 59–78.
- Byrne, D. E. (1971). *The attraction paradigm* (Vol. 462). Academic Press.
- Casaló, L., Flavián, C., & Guinalfú, M. (2007). The impact of participation in virtual brand communities on consumer trust and loyalty. *Online Information Review*, 31(6), 775–792.
- Chattalas, M., & Takada, H. (2013). Warm versus competent countries: National stereotyping effects on expectations of hedonic versus utilitarian product properties. *Place Branding and Public Diplomacy*, 9(2), 88–97.
- Chaudhuri, A., & Holbrook, M. B. (2001). The chain of effects from brand trust and brand affect to brand performance: The role of brand loyalty. *Journal of Marketing*, 65(2), 81–93.
- 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.
- Cohrs, J. C., Asbrock, F., & Sibley, C. G. (2012). Friend or foe, champ or chump? Social conformity and superiority goals activate warmth-versus competence-based social categorization schemas. *Social Psychological and Personality Science*, 3(4), 471–478.
- Crowley, A. E., Spangenberg, E. R., & Hughes, K. R. (1992). Measuring the hedonic and utilitarian dimensions of attitudes toward product categories. *Marketing Letters*, 3(3), 239–249.
- Cuddy, A. J., Fiske, S. T., & Glick, P. (2008). Warmth and competence as universal dimensions of social perception: The stereotype content model and the BIAS map. *Advances in Experimental Social Psychology*, 40, 61–149.
- Dahlbäck, N., Jönsson, A., & Ahrenberg, L. (1993). Wizard of Oz studies: Why and how. In *Proceedings of the 1st international conference on Intelligent user interfaces* (pp. 193–200).
- DeArmond, S., Tye, M., Chen, P. Y., Krauss, A., Apryl Rogers, D., & Sintek, E. (2006). Age and gender stereotypes: New challenges in a changing workplace and workforce. *Journal of Applied Social Psychology*, 36(9), 2184–2214.

- Dhar, R., & Wertenbroch, K. (2000). Consumer choice between hedonic and utilitarian goods. *Journal of Marketing Research*, 37(1), 60–71.
- Dijksterhuis, A., & Bargh, J. A. (2001). The perception-behavior expressway: Automatic effects of social perception on social behavior. In *Advances in experimental social psychology* (Vol. 33, pp. 1–40). Academic Press.
- Duffy, B. R. (2003). Anthropomorphism and the social robot. *Robotics and Autonomous Systems*, 42(3–4), 177–190.
- Eagly, A. H., & Steffen, V. J. (1984). Gender stereotypes stem from the distribution of women and men into social roles. *Journal of Personality and Social Psychology*, 46(4), 735–754.
- Eagly, A. H. (2013). *Sex differences in social behavior: A social-role interpretation*. Psychology Press.
- Edwards, C., Edwards, A., Stoll, B., Lin, X., & Massey, N. (2019). Evaluations of an artificial intelligence instructor's voice: Social Identity Theory in human-robot interactions. *Computers in Human Behavior*, 90, 357–362.
- Epley, N., Waytz, A., & Cacioppo, J. T. (2007). On seeing human: A three-factor theory of anthropomorphism. *Psychological Review*, 114(4), 864–886.
- Eyssel, F., & Hegel, F. (2012). (S) he's got the look: Gender stereotyping of robots. *Journal of Applied Social Psychology*, 42(9), 2213–2230.
- Fiske, S. T., Cuddy, A. J. C., Glick, P., & Xu, J. (2002). A model of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and competition. *Journal of Personality and Social Psychology*, 82(6), 878–902.
- Fiske, S. T., Cuddy, A. J. C., & Glick, P. (2007). Universal dimensions of social cognition: Warmth and competence. *Trends in Cognitive Sciences*, 11(2), 77–83.
- Fiore, S. M., Wiltshire, T. J., Lobato, E. J., Jentsch, F. G., Huang, W. H., & Axelrod, B. (2013). Toward understanding social cues and signals in human-robot interaction: Effects of robot gaze and proxemic behavior. *Frontiers in Psychology*, 4, 859.
- Fournier, S., & Alvarez, C. (2012). Brands as relationship partners: Warmth, competence, and in-between. *Journal of Consumer Psychology*, 22(2), 177–185.
- Gao, Y. L., & Mattila, A. S. (2014). Improving consumer satisfaction in green hotels: The roles of perceived warmth, perceived competence, and CSR motive. *International Journal of Hospitality Management*, 42, 20–31.
- Gupta, V. K., Turban, D. B., & Bhawne, N. M. (2008). The effect of gender stereotype activation on entrepreneurial intentions. *Journal of Applied Psychology*, 93(5), 1053–1061.
- Gentsch, P. (2018). *AI in marketing, sales and service: How marketers without a data science degree can use AI, big data and bots*. Springer.
- Hamilton, D. L., Sherman, S. J., & Ruvalo, C. M. (1990). Stereotype-based expectancies: Effects on information processing and social behavior. *Journal of Social Issues*, 46(2), 35–60.
- Hayes, A. F. (2017). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. Guilford Publications.
- Hayes, A. F., & Scharkow, M. (2013). The relative trustworthiness of inferential tests of the indirect effect in statistical mediation analysis: Does method really matter? *Psychological Science*, 24(10), 1918–1927.
- Heilman, M. E. (2012). Gender stereotypes and workplace bias. *Research in Organizational Behavior*, 32, 113–135.
- Holbrook, M. B., & Hirschman, E. C. (1982). The experiential aspects of consumption: Consumer fantasies, feelings, and fun. *Journal of Consumer Research*, 9(2), 132. <https://doi.org/10.1086/jcr.1982.9.issue-210.1086/208906>
- Huang, C. (2020). Smart speaker ads on the increase in China. *Campaign*. <https://www.campaignasia.com/article/smart-speaker-ads-on-the-increase-in-china/461599> (accessed June 30, 2021).
- Huddy, L., & Terkildsen, N. (1993). Gender stereotypes and the perception of male and female candidates. *American Journal of Political Science*, 37(1), 119. <https://doi.org/10.2307/2111526>
- Insider Intelligence (2021). Chatbot market in 2021: Stats, trends, and companies in the growing AI chatbot industry. *Insider*. <https://www.businessinsider.com/chatbot-market-stats-trends> (accessed June 30, 2021).
- Kornadt, A. E. (2016). Do age stereotypes as social role expectations for older adults influence personality development? *Journal of Research in Personality*, 60, 51–55.
- Kim, A., Cho, M., Ahn, J., & Sung, Y. (2019). Effects of gender and relationship type on the response to artificial intelligence. *Cyberpsychology, Behavior, and Social Networking*, 22(4), 249–253.
- Kolbl, Z., Arslanagic-Kalajdzic, M., & Diamantopoulos, A. (2019). Stereotyping global brands: Is warmth more important than competence? *Journal of Business Research*, 104, 614–621.
- Lee, K.-M., & Nass, C. (2005). Social-psychological origins of feelings of presence: Creating social presence with machine-generated voices. *Media Psychology*, 7(1), 31–45.
- Malone, C., & Fiske, S. T. (2013). *The human brand: How we relate to people, products, and companies*. John Wiley & Sons.
- Moon, Y. (2000). Intimate exchanges: Using computers to elicit self-disclosure from consumers. *Journal of Consumer Research*, 26(4), 323–339.
- Nass, C., Steuer, J., & Tauber, E. R. (1994). Computers are social actors. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 72–78).
- Nass, C., Moon, Y., Fogg, B. J., Reeves, B., & Dryer, C. (1995). Can computer personalities be human personalities? In *Conference companion on Human factors in computing systems* (pp. 228–229).
- Nass, C., Moon, Y., & Green, N. (1997). Are machines gender neutral? Gender-stereotypic responses to computers with voices. *Journal of Applied Social Psychology*, 27(10), 864–876.
- Nass, C., & Moon, Y. (2000). Machines and mindlessness: Social responses to computers. *Journal of Social Issues*, 56(1), 81–103.
- Park, C. (2004). Efficient or enjoyable? Consumer values of eating-out and fast food restaurant consumption in Korea. *International Journal of Hospitality Management*, 23(1), 87–94.
- Parulekar, A. A., & Raheja, P. (2006). Managing celebrities as brands: Impact of endorsements on celebrity image. In *Creating images and the psychology of marketing communication* (pp. 161–169).
- 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.
- Pratt, J. A., Hauser, K., Ugray, Z., & Patterson, O. (2007). Looking at human-computer interface design: Effects of ethnicity in computer agents. *Interacting with Computers*, 19(4), 512–523.
- Preacher, K. J., Rucker, D. D., & Hayes, A. F. (2007). Addressing moderated mediation hypotheses: Theory, methods, and prescriptions. *Multivariate Behavioral Research*, 42(1), 185–227.
- Reeves, B., & Nass, C. (1996). *The media equation: How people treat computers, television, and new media like real people*. Cambridge, United Kingdom: Cambridge University Press.
- Rhee, C. E., & Choi, J. (2020). Effects of personalization and social role in voice shopping: An experimental study on product recommendation by a conversational voice agent. *Computers in Human Behavior*, 109, 106359. <https://doi.org/10.1016/j.chb.2020.106359>
- Roy, R., & Naidoo, V. (2021). Enhancing chatbot effectiveness: The role of anthropomorphic conversational styles and time orientation. *Journal of Business Research*, 126, 23–34.
- Scarpi, D. (2021). A construal-level approach to hedonic and utilitarian shopping orientation. *Marketing Letters*, 32(2), 261–271.
- Sung, Y., & Kim, J. (2010). Effects of brand personality on brand trust and brand affect. *Psychology & Marketing*, 27(7), 639–661.
- Tay, B. T. C., Park, T., Jung, Y., Tan, Y. K., & Wong, A. H. Y. (2013). When stereotypes meet robots: The effect of gender stereotypes on people's acceptance of a security robot. In *International conference on engineering psychology and cognitive ergonomics* (pp. 261–270). Berlin, Heidelberg: Springer.
- Vogel, D. L., Wester, S. R., Heesacker, M., & Madon, S. (2003). Confirming gender stereotypes: A social role perspective. *Sex Roles*, 48(11–12), 519–528.
- Weizenbaum, J. (1976). *Computer power and human reason: From judgment to calculation*. Williams, J. E., & Best, D. L. (1990). *Measuring sex stereotypes: A multinational study*. Rev. Sage Publications, Inc.
- World Economic Forum (2019). Global gender gap report 2020. (Retrieved December 26, 2020, from) [http://www3.weforum.org/docs/WEF\\_GGGR\\_2020.pdf/](http://www3.weforum.org/docs/WEF_GGGR_2020.pdf/).
- Yeung, C. W., & Wyer, R. S., Jr. (2004). Affect, appraisal, and consumer judgment. *Journal of Consumer Research*, 31(2), 412–424.
- Zhu, D. H., & Deng, Z. Z. (2021). Effect of social anxiety on the adoption of robotic training partner. *Cyberpsychology, Behavior, and Social Networking*, 24(5), 343–348.
- Zuckerman, M., & Driver, R. E. (1988). What sounds beautiful is good: The vocal attractiveness stereotype. *Journal of Nonverbal Behavior*, 13(2), 67–82.

**Jungyong Ahn** is a research professor at the School of Media & Communication, Korea University. His research interest focuses on brand management, consumer-brand relationships, new media advertising, artificial intelligence in consumer behavior.

**Jungwon Kim** is a doctoral student at the School of Psychology, Korea University. Her research focuses on consumer and advertising psychology, especially usage of new media, interactions between AI and human, and privacy in digital media.

**Yongjun Sung** is a professor of psychology at the School of Psychology, Korea University. His research focuses on self-concept, brand personality, brand management, digital media, and artificial intelligence (AI). He has authored or co-authored over 100 articles in leading referred journals including *Journal of Business Research*, *Journal of Consumer Psychology*, *Journal of Advertising*, *Psychology & Marketing*, *Journal of Advertising*, *International Journal of Advertising*, among others.