Computers in Human Behavior 26 (2010) 665–672

Contents lists available at ScienceDirect

Computers in Human Behavior

journal homepage: www.elsevier.com/locate/comphumbeh

The more humanlike, the better? How speech type and users' cognitive style affect social responses to computers

Eun-Ju Lee *

Department of Communication, Seoul National University, San 56-1 Shilim-Dong, Gwanak-Gu, Seoul 151-742, South Korea

A R T I C L E I N F O

Article history:
Available online 29 January 2010

Keywords:
Anthropomorphism
Computers Are Social Actors (CASA)
Experientiality
Rationality

A B S T R A C T

The present experiment investigated if anthropomorphic interfaces facilitate people's tendency to project social expectations onto computers and how such effects might vary depending on users' cognitive style. In a 2 (synthetic vs. recorded speech) × 2 (flattering vs. generic feedback) × 2 (low vs. high rationality) × 2 (low vs. high experientiality) experiment, participants played a trivia game with a computer. Use of recorded speech did not amplify the previously documented flattery effects (Fogg & Nass, 1997), challenging the notion that anthropomorphism will promote social responses to computers. Participants evaluated the human-voiced computer more positively and conformed more to its suggestions than the one using synthetic speech, but such effects were found only among less analytical or more intuition-driven individuals, suggesting dispositional differences in people's susceptibility to anthropomorphic cues embedded in the interface.

1. Introduction

Although it might sound unreasonable, the notion that people treat computers as if they were humans (Nass & Brave, 2005; Nass & Moon, 2000; Reeves & Nass, 1996) has received consistent empirical support using a wide array of social rules observed in human–human relationships. According to the research paradigm called "Computers Are Social Actors" (CASA), despite their knowledge that computers do not warrant social treatment, people nonetheless tend to apply social expectations to and exhibit the same responses to computers as they would to human interactants. For example, consistent with gender stereotypes associating each gender with different domains of expertise (e.g., Carli, 2001), when the computer gender was manifested in gender-marked cartoon characters, participants were more likely to accept the male than female computer's suggestions for sports questions, and vice versa for fashion trivia (Lee, 2003). Likewise, just as people are more attracted to similar than dissimilar others (Byrne, 1971), when the computer's 'personality' was implicitly conveyed through textual cues, participants responded more favorably to the computer whose personality matched their own (Moon & Nass, 1996); that is, dominant users preferred a dominant computer that displayed greater self-confidence and used more assertive language, whereas submissive users found the interaction more satisfying when the computer used tentative language expressing lower self-confidence.

Even when a computer does not embody such uniquely human traits, like gender and personality, people seem to spontaneously follow social rules in their dealings with computers. For instance, when the computer had first provided some information about its technical capacity, participants reciprocated such self-disclosure by providing more intimate responses to the computer's questions (Moon, 2000). In a similar vein, Fogg and Nass (1997) found that flattery from a computer had the same positive effects as that from another person. Even though participants were explicitly informed that the computer's evaluations were generated randomly, and thus, not diagnostic of the actual quality of their performance, such meaningless praises from the computer led them to think that they had indeed performed better and evaluate the computer more favorably than when the computer offered generic feedback.

After reviewing these findings demonstrating the robust operation of social rules in human–computer interaction (HCI), it might seem only natural to make the interface more humanlike, thereby conforming to the prevalent social expectations of users. Put it differently, if people are generally predisposed to treat computers as if they were humans, it might be all the better to have them look and act like one, and thus, require no adjustments on the part of users (Reeves & Nass, 1996). However, empirical research has failed to yield consistent support for this seemingly intuitive notion (see Dehn & van Mulken, 2000 for a review), prompting the question of how humanlike attributes of the interface alter the ways in which people interact with a computer. For example, when do anthropomorphic interfaces elicit more positive or

* Tel.: +82 2 880 6469; fax: +82 2 885 8418.
E-mail address: eunju0204@snu.ac.kr

0747-5632/$ - see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.chb.2010.01.003

negative reactions from the users than their mechanomorphic counterparts? Are there any individual differences in how people respond to the more or less humanlike interfaces? Moreover, with respect to CASA, will various cues to humanity manifested in the interface, such as human voice and anthropomorphic agents, facilitate people's proclivity to apply social rules to the computer?

The present research aimed to address these questions by focusing on (a) how human likeness of the interface influences the extent to which people project social expectations onto the computer and (b) how users' cognitive style moderates the ways in which they respond to the anthropomorphic interface. Specifically, using the aforementioned flattery effects (Fogg & Nass, 1997) as an example of social responses, the current experiment first investigated if the use of recorded (vs. synthesized) speech output would amplify the previously documented flattery effects. Second, considering that the anthropomorphic feature implemented herein represented only the surface-level variation, affecting the looks and feels of the computer system with no substantial functional improvement, it was predicted that users' cognitive propensity might moderate the relative (dis)advantages of the anthropomorphic interface. Specifically, based on the cognitive-experiential self-theory (CEST; Epstein, 1994; Epstein & Pacini, 1999), the current research examined if those less prone to analytical thinking (low rationals) or those relying more heavily on heuristic principles (high experientials) would show greater sensitivity to anthropomorphic features of the interface and respond differently to more or less human-sounding computers.

2. The more humanlike, the more social responses to computers?

Although the tendency to anthropomorphize computers is considered to be rather universal in the CASA tradition, some researchers have examined how various anthropomorphic qualities of the interface affect user responses. Sproull, Subramani, Kiesler, and Walker (1996), for example, used an ostensible computer-based career counseling system, which offered its output through either simple text display or a talking-face. As expected, participants reported greater arousal and were more likely to engage in impression management in the talking-face condition, presumably because "the talking-face display reminded subjects of a real human being (p. 117)". In a follow-up study, participants played a social dilemma game with a human confederate via videoconferencing or with one of three interface agents: a person-like agent, a dog-like agent, or a cartoon dog agent. Results showed that the humanlike agent elicited greater cooperation, to the extent comparable to the human confederate, although the dog-like agent was perceived to be more likable (Parise, Kiesler, Sproull, & Waters, 1999). Extending this appearance-based approach to human–robot interaction, Hinds, Roberts, and Jones (2004) examined how the humanlike appearance of professional service robots (e.g., facial features, arms and legs, human clothing) influences the working relationship, in conjunction with the status manipulation (subordinate vs. peer vs. supervisor). They found that participants took more credit for the successful completion of the task when working with a machine-like robot, especially the one assigned the subordinate status, which suggests that anthropomorphic appearance helps to establish equal partnership in human–robot collaboration.

Moving beyond a comparison between more or less human-looking interfaces, other researchers have created more subtle variations in the computer agent's nonverbal behaviors and investigated how such changes affect interaction patterns. Specifically, Bailenson, Blascovich, Beall, and Loomis (2001) examined how embodied computer agents' gaze affects the interpersonal distance people maintain in the immersive virtual reality environment. When the virtual human displayed realistic gaze, thereby more closely approximating human–human interaction, people increased interpersonal distance to maintain an optimal level of immediacy. Similarly, Cassell and Thorisson (1999) compared two different kinds of nonverbal behaviors of embodied computer agents: envelope feedback (i.e., those related to the process of conversation, such as nods, gaze, and head movements) and emotional feedback (i.e., facial displays of such emotions as happiness and puzzlement). They found that when the agent displayed the envelope feedback, which represents a more sophisticated form of nonverbal communication conducive to smoother social transaction, participants not only evaluated the interaction more positively, but they also managed to use fewer utterances to complete the task.

Albeit informative, previous studies focused mostly on how the anthropomorphic interfaces compare to their non- or less-anthropomorphic counterparts, in terms of users' liking, communication behavior, and task performance. Consequently, it remains to be answered how human likeness of the interface might affect the propensity to treat computers as social actors. Although Sundar (2004) measured individuals' overall tendency to anthropomorphize computers and examined how it relates to their loyalty to a particular terminal in public computer clusters, his study concerned dispositional differences in the extent to which individuals consciously make social attributions to computers, as opposed to the situational triggers of social responses. To fill this void, the current research systematically varied the degree of humanness manifested in the computer output (i.e., recorded human speech vs. synthesized speech) and examined how it moderates the likelihood of social responses to computers that flatter.

In fact, several CASA studies used synthetic speech and still found social responses to computers (e.g., Lee & Nass, 2004; Lee, Nass, & Brave, 2000), indicating the robust tendency to humanize computers. More recent findings, however, have suggested that clearly machine-like speech might inhibit, if not completely suppress, social responses to computers. For example, whereas the recorded voice system was perceived as more useful when saying "I" (e.g., I will begin today's auction shortly), the synthetic voice system was rated as less trustworthy and less useful with "I" (Nass & Brave, 2005, p. 119), which indicates that people are reluctant to assign full humanity to machine-generated voices. In addition, although gender perception was more salient when the computer produced its output in synthetic speech than in written text with gendered cartoon characters, participants did not exhibit gender stereotyping toward the computer using synthesized voices (Lee, 2008a). Taken together, these findings seem to suggest that computer-generated speech might serve as a cue to remind people of the *a*social nature of the interaction, thereby hindering social responses. By directly comparing synthetic and recorded human voices, the current research aimed to evaluate this conjecture. The following hypotheses were drawn in line with the original flattery study with an emphasis on the moderating role of the voice type.

H1a–b. Participants will perceive the flattering computer as (a) more socially attractive and (b) more believable than the one producing generic feedback. However, such effects will be more pronounced when the computer uses recorded human speech rather than synthesized speech.

H2a–b. Participants will evaluate (a) their own performance and (b) the computer's overall performance more positively when the computer provides flattering than generic feedback. However, such effects will be more pronounced when the computer uses recorded human speech rather than synthesized speech.

Unlike the original study that focused on affective reactions to the flattering computer, the present research also measured behavioral responses in terms of conformity. Specifically, participants played a trivia game with a computer. On critical trials, the computer claimed that the participant had picked a wrong answer and suggested a different one. Knowing that the computer was programmed to offer random feedback, participants were then given a chance to submit their final answer, which would determine their performance score. Such a decision posed a dilemma by forcing participants to decide either sticking to their initial pick or shifting to the computer's suggestion. Possibly, when the computer praised their performance on previous trials, flattery-induced positive feelings about the computer might make participants *more* willing to accept the computer's answers, as liking increases persuasive power of the influence agent (Eagly & Chaiken, 1975). Conversely, if positive feedback from the computer boosts the participants' confidence about their own performance (Fogg & Nass, 1997), it might make them *less* susceptible to any influence attempts. In fact, one study found that flattery lowered conformity to the computer, as compared to strictly factual feedback, although it was unclear why such was the case (Lee, 2008b). Given these competing possibilities, the following research question was proposed in lieu of a formal hypothesis.

RQ1: How will flattery affect the acceptance of the computer's suggestions? Will the effects vary depending on the speech type (recorded vs. synthetic speech)?

3. Anthropomorphism as a peripheral cue: Interaction with users' cognitive style

Despite wide disparities in how anthropomorphic interface was operationalized in previous research, ranging from the display of a talking-face (Sproull et al., 1996) to nonverbal feedback (Cassell & Thorisson, 1999) to eye gaze (Bailenson et al., 2001), it is noteworthy that human likeness, in and of itself, does not represent fundamental improvement in the computer's functionality. More often than not, it instead serves as a peripheral cue that creates potentially misleading impressions about its technical capacity. Consequently, some researchers have warned of the possibility that overly anthropomorphic agents might backfire when their behaviors fail to map onto users' expectations (Bailenson et al., 2005; Heckman & Wobbrock, 2000). That is, users might come to liken the computer to a human on the basis of its surface-level resemblance and expect it to fully function like a real person, which in turn, almost inevitably leads to disappointment and frustration. In fact, such potentially negative outcome is well captured in the notion of "uncanny valley" (Mori, 1970), which refers to the user's revulsion to subtle deviations from human norms manifested by near-humanlike forms, such as androids. Although more recent studies have shown that human likeness is not the only determinant of perceived eeriness or strangeness of a robot or animated characters, and thus, the uncanny valley can be escaped by carefully implementing other factors into the design (e.g., MacDorman, 2006), it is widely accepted that human likeness of computer agents plays a significant role in forming users' perceptions of and reactions to them, for better or for worse.

If human likeness serves as a peripheral cue that shapes users' expectations, its effects might vary depending on user characteristics. According to cognitive-experiential self-theory (CEST), there are two different information-processing systems, experiential and rational (Epstein, 1994; Epstein & Pacini, 1999). Just as systematic or central processing entails a relatively analytic and comprehensive treatment of judgment-relevant information (Chen & Chaiken, 1999; Petty & Cacioppo, 1986), the rational system involves highly effortful cognitive activities. On the other hand, the experiential system resembles heuristic or peripheral processing as it operates in a rather automatic and holistic manner, demanding minimal cognitive resources (Epstein & Pacini, 1999). Unlike other dual-process models, however, CEST assumes that rational and experiential thinking styles are enduring dispositions that predispose individuals to heed and process information differently. Given that peripheral or heuristic cues exert greater influence on attitude change when message recipients engage in less effortful and analytical thinking (Chen & Chaiken, 1999; Petty & Cacioppo, 1986), merely adding a human voice in lieu of machine-generated one, with no corresponding functional changes, might also have greater impact on low rationals, who are chronically less prone to effortful cognitive activities, or high experientials, who rely more heavily on peripheral cues. By examining individual differences in the susceptibility to nominal anthropomorphic cues, the present research aimed to specify the contingent conditions under which anthropomorphic qualities of the interface are more or less likely to affect user experience.

Predicting the effects of anthropomorphic interface is not straightforward, though. Previous research has indicated that human likeness can have either positive or negative effects depending on the domain. For example, participants were less willing to perform embarrassing acts (Bailenson et al., 2005) and displayed a stronger social desirability bias in front of more human-looking agents (Sproull et al., 1996), which suggests that when the task involves socially sensitive topics, users might be more comfortable with relatively mechanomorphic interfaces. Based on a meta-analysis, Dehn and van Mulken (2000) concluded that "the empirical evidence on the way an animated anthropomorphized agent influences the user's attitudes towards the system (e.g., utility, likeability, and comfortability) appears to be mixed" except that "a system with an agent is perceived as more entertaining than the one without an agent" (p. 15). They also pointed out that when facial display attracts users' attention, it might hinder task performance. However, in a more recent meta-analysis, Yee, Bailenson, and Rickertsen (2007) reviewed studies that compared interfaces with visually embodied agents with those without agents, and found that facial representation yielded more positive social interactions, especially in terms of subjective responses. Therefore, a research question was proposed as to if more or less human-sounding computers elicit different affective and behavioral responses, and if so, such effects vary as a function of the user's cognitive propensity.

RQ2: How will the user's cognitive style moderate the effects of the speech type (recorded vs. synthesized speech) on (a) participants' evaluations of the computer and (b) their acceptance of the computer's suggestions?

4. Method

4.1. Participants

Participants were 187 college undergraduates (68 men, 119 women) enrolled in communication classes at a large west-coast university in the United States. They took part in a 2 (Computer Comments: Flattery vs. Generic Comments) $\times$ 2 (Speech Type: Recorded vs. Synthesized Speech) $\times$ 2 (Dichotomized Rationality: High vs. Low) $\times$ 2 (Dichotomized Experientiality: High vs. Low) between-subjects design experiment.

4.2. Procedure

Upon arrival at the laboratory, participants were told that they would play an interactive trivia game with a computer. To enhance a sense of interactivity, they were first asked to choose a number, ranging from 1 to 10, to determine a set of questions. In actuality, the computer presented the same questions irrespective of the number selected. The computer then presented a multiple-choice trivia question (e.g., "Which of the following famous couples had the shortest marriage?"). Once the participants picked their initial answer and moved to the next page, the computer gave its answer in either synthetic or human voice (e.g., "Sorry, the correct answer is D. Drew Barrymore and Jeremy Thomas were married for two weeks"). On five critical trials, the computer presented a different answer from the participant's. Because participants were told that the computer was programmed to generate random answers, their main task was to guess whether the computer was presenting the correct answer or not, and submit their final answer accordingly. In order to control for the effects of computer gender, two male voices were used for each speech type.

On seven filler trials, the computer validated participants' initial answer, providing either generic or flattering comments. Generic comments had neither positive nor negative valence, but simply conveyed factual information (e.g., "That's right. Spain gave women the right to vote in 1893"). In addition to this factual feedback, those in the flattery condition received various praising remarks, such as "Excellent job!" or "Your trivia knowledge is quite exceptional!"

To enhance their involvement in the experiment and discourage blind rejection of the computer's answer, a $20 cash prize was offered for the person with the highest score. Thus, to maximize the likelihood of winning the prize, participants had to accept the computer's answer when it seemed like a correct one. Once they submitted their final answer, the computer proceeded to the next question without revealing the real answer. This procedure was repeated for 12 different questions.

4.3. Index construction

For social attractiveness and believability, participants were presented with a list of adjectives on a post-test questionnaire. They indicated how well each word described the computer they interacted with on a 10-point scale (1 – *describes very poorly*, 10 – *describes very well*). A factor analysis yielded a two-factor solution that accounted for 74.74% of the original variance. The first factor, which represented social attractiveness, consisted of *likable*, *sociable*, *friendly*, and *personal* (Eigenvalue = 3.68, $\alpha = .84$). The ratings were then summed across these four items ($M = 20.82$, $SD = 7.46$). The second factor captured believability, comprised of the remaining three items: *trustworthy*, *reliable*, *honest* (Eigenvalue = 1.55, $\alpha = .90$, $M = 12.50$, $SD = 5.76$).

Participants' self-evaluation was measured by asking for how many questions they thought they answered correctly ($M = 4.34$, $SD = 1.98$). For overall evaluation of the computer performance, they also estimated for how many questions they thought the computer had provided the correct answer ($M = 5.57$, $SD = 2.08$).

Conformity was measured by counting how many times, out of the five critical questions for which the computer suggested a different answer, participants switched to the computer's suggestion. Conformity score ranged from 0 to 5 ($M = 2.38$, $SD = 1.51$). After choosing the final answer for each question, participants also estimated how likely the computer had presented the correct answer on an 11-point scale, ranging from 0% to 100% in 10% increments. The ratings were then averaged to yield the index for perceived validity of the computer feedback ($\alpha = .86$, $M = 40.95$, $SD = 20.88$).

Participants' cognitive style was measured by the short version of Rational–Experiential Inventory (REI) (Pacini & Epstein, 1999). Items for rationality captured both self-reported ability to think logically and analytically (e.g., "I am much better at figuring things out logically than most people") and enjoyment of thinking in an analytical manner (e.g., "I prefer complex to simple problems"). Similarly, experientiality was measured by items capturing both preference and self-assessed ability for intuitive judgments, such as "I often go by my instincts when deciding on a course of action" and "Using my gut feelings usually works well for me in figuring out problems in my life". Responses were made on a five-point scale ranging from *Definitely False* (1), *Mostly False* (2), *Undecided or Equally True and False* (3), *Mostly True* (4), to *Definitely True* (5), and summed across 12 items to create the rationality and experientiality indices, respectively. Rationality scores ranged from 22 to 59 ($\alpha = .87$, $M = 42.53$, $SD = 6.93$) and the experientiality scores ranged from 25 to 56 ($\alpha = .84$, $M = 41.97$, $SD = 6.03$). They were not significantly correlated with each other ($r = .03$), confirming their orthogonality, as suggested by CEST. The cutoff for the median split was 43 for both rationality and experientiality indices.

5. Results

5.1. Manipulation check

To establish if the use of recorded speech rendered the computer more humanlike, as compared to the synthetic speech, participants were asked to rate the computer they interacted with on 10-point semantic differential scales: unnatural/natural, machine-like/humanlike, artificial/lifelike (Powers & Kiesler, 2006). The ratings were then averaged to yield the perceived anthropomorphism score ($\alpha = .65$, $M = 5.40$, $SD = 1.77$). An independent-samples *t*-test confirmed that participants perceived the computer using recorded speech ($M = 5.78$, $SD = 1.92$) as more anthropomorphic than the synthetic-voice one ($M = 4.99$, $SD = 1.50$), $t(185) = 3.09$, $p = .002$, Cohen's $d = .46$. Although the difference between the two speech types was statistically significant, indicating that the induction of human likeness was effective, it seems noteworthy that its magnitude was relatively small.

5.2. Hypothesis tests

H1a–b predicted that recorded speech will amplify the positive effects of flattery on perceived social attractiveness and believability. First, a $2 \times 2 \times 2 \times 2$ (Computer Comments $\times$ Speech Type $\times$ Dichotomized Rationality $\times$ Dichotomized Experientiality) ANOVA was computed on perceived social attractiveness. A significant interaction emerged between computer comments and speech type, $F(1, 171) = 5.20$, $p = .02$, $\eta_p^2 = .03$ (see Table 1 for descriptive statistics). Simple effects tests showed that, counter to *H1a*, flattery increased perceived social attractiveness, as compared to generic comments, only when delivered in synthetic speech, $F(1, 183) = 10.09$, $p = .002$, $\eta_p^2 = .05$. When recorded speech was used, there was no significant effect of computer comments, $F < 1$. The significant main effect for computer comments, with

Table 1
Means and standard deviations for social attractiveness and believability.

	TTS		Recorded Speech	
	Flattery	Generic comments	Flattery	Generic comments
Social Attractiveness	23.82 (8.75)	18.93 (6.35)	20.34 (7.29)	20.26 (6.64)
Believability	13.24 (6.31)	11.00 (5.58)	11.76 (5.00)	13.88 (5.82)

flattery ($M = 22.04$, $SD = 8.18$) garnering higher social attractiveness ratings than generic comments ($M = 19.63$, $SD = 6.50$), should be interpreted in the context of this disordinal interaction, $F(1, 171) = 4.75$, $p = .03$, $\eta_p^2 = .03$.

A $2 \times 2 \times 2 \times 2$ ANOVA on perceived believability of the computer also yielded a significant computer comments by speech type interaction, $F(1, 171) = 6.80$, $p = .01$, $\eta_p^2 = .03$, but the pattern was slightly different (see Table 1). Specifically, when the computer used synthetic speech, participants tended to rate the flattering computer as *more* believable than the generic-comment computer, $F(1, 183) = 3.50$, $p = .06$, $\eta_p^2 = .02$. By contrast, when recorded speech was used, flattering computer was considered as somewhat *less* believable than its generic-comment counterpart, $F(1, 183) = 3.41$, $p = .07$, $\eta_p^2 = .02$. There were no other significant main or interaction effects, all Fs < 2.22, all ps > .13, indicating no moderating role of users' cognitive style (*RQ2*).

H2a–b addressed the questions of how flattery from the computer affects the participants' self-evaluation and the overall evaluation of the computer's performance, especially in conjunction with the speech type. A $2 \times 2 \times 2 \times 2$ (Computer Comments $\times$ Speech Type $\times$ Dichotomized Rationality $\times$ Dichotomized Experientiality) ANOVA on the participants' estimate of the number of correct answers yielded no significant interaction between computer comments and speech type, failing to support *H2a*. Instead, a significant computer comments by dichotomized rationality interaction emerged, $F(1, 171) = 11.41$, $p = .001$, $\eta_p^2 = .06$ (see Fig. 1). Specifically, low rationals thought they had done a better job when flattered ($M = 4.54$, $SD = 1.72$) than when given generic feedback ($M = 3.64$, $SD = 1.87$), $F(1, 183) = 5.59$, $p = .02$, $\eta_p^2 = .03$. By contrast, high rationals became less confident about their own performance given bogus praise ($M = 4.18$, $SD = 1.96$) than factual feedback ($M = 5.11$, $SD = 2.09$), $F(1, 183) = 5.28$, $p = .02$, $\eta_p^2 = .03$. When the interaction was decomposed within each comment type, high rationals exhibited greater self-confidence than their less analytical counterparts when generic feedback was given, $F(1, 183) = 13.93$, $p < .001$, $\eta_p^2 = .07$, but flattering feedback from the computer removed such difference between high and low rationals, $F < 1$. Likewise, there was a significant interaction between user's experiential thinking style and computer comments (*RQ2*), $F(1, 171) = 3.95$, $p < .05$, $\eta_p^2 = .02$; that is, those favoring intuitive thinking and trusting their gut feelings assessed their self-performance more positively ($M = 4.72$, $SD = 1.90$) than did their less intuition-relying counterparts ($M = 3.82$, $SD = 1.68$), but such difference was observed only when the computer offered flattering comments, $F(1, 183) = 4.62$, $p = .04$, $\eta_p^2 = .02$. When generic feedback was presented, high ($M = 4.27$, $SD = 1.86$) and low ($M = 4.39$, $SD = 2.28$) experientials were not significantly different in their self-evaluation, $F < 1$.

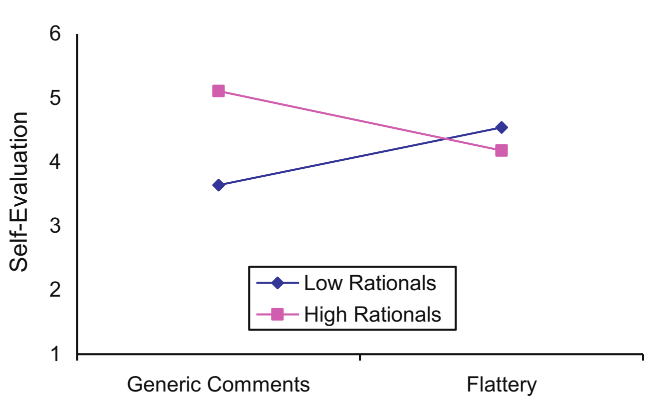

Fig. 1. Self-evaluation by rationality and computer comments.

With respect to the evaluation of the computer performance, there were no significant main or interaction effects involving computer comments, failing to support *H2b*, all Fs < 2.30, all ps > .13. However, as suspected in *RQ2*, speech type had significant interactions with both participants' rationality, $F(1, 171) = 8.40$, $p = .004$, $\eta_p^2 = .05$, and experientiality, $F(1, 171) = 6.34$, $p = .01$, $\eta_p^2 = .04$ (see Table 2 for descriptive statistics). First, low rationals rated the human-sounding computer more positively than the machine-like one, $F(1, 183) = 6.14$, $p = .01$, $\eta_p^2 = .03$, whereas high rationals did not discriminate the computer using synthetic voice, $F(1, 183) = 2.59$, $p = .11$. Just as low rationals, high experientials perceived the human-voiced computer to have outperformed its mechanic-voiced counterpart, $F(1, 183) = 4.12$, $p = .04$, $\eta_p^2 = .02$, with those less trusting their intuition making no corresponding differentiation, $F(1, 183) = 1.51$, $p = .22$.

RQ1 concerned if speech type moderates flattery effects on conformity. Flattery tended to elicit less conformity ($M = 2.14$, $SD = 1.49$) than did generic comments ($M = 2.61$, $SD = 1.51$), but the difference failed to reach statistical significance, $F(1, 171) = 3.29$, $p = .07$, $\eta_p^2 = .02$. Moreover, speech type had no significant interaction with computer comments, $F < 1$. Instead, a significant interaction emerged between dichotomized rationality and speech type (*RQ2*), $F(1, 171) = 5.43$, $p = .02$, $\eta_p^2 = .03$ (see Fig. 2). Although low rationals were more likely to accept the computer's suggestions when it used recorded human speech ($M = 2.83$, $SD = 1.36$), as opposed to synthetic speech ($M = 2.19$, $SD = 1.61$), $F(1, 183) = 4.44$, $p = .04$, $\eta_p^2 = .02$, high rationals showed no such differentiation, $F < 1$. In addition, high experientials ($M = 2.14$, $SD = 1.48$) were less likely to switch to the computer's recommendation than were lows ($M = 2.62$, $SD = 1.52$), presumably because they trusted their gut feelings and thus, were less willing to give up their initial pick, $F(1, 171) = 4.45$, $p = .04$, $\eta_p^2 = .03$.

Likewise, a $2 \times 2 \times 2 \times 2$ ANOVA on perceived validity of the computer's suggestion, which was measured for each question after the participants chose their final answer, yielded a significant experientiality by speech type interaction (*RQ2*), $F(1, 171) = 5.24$, $p = .02$, $\eta_p^2 = .03$. When synthetic voice informed the participants that they had picked a wrong answer, high experientials ($M = 37.50$, $SD = 17.77$) expressed greater suspicion about the computer's claim than did low experientials ($M = 47.52$, $SD = 22.55$), $F(1, 183) = 5.12$, $p = .03$, $\eta_p^2 = .03$. With human voice delivering the verdict, however, high ($M = 42.78$, $SD = 20.92$) and low experientials ($M = 37.14$, $SD = 21.17$) showed no significant differences in their assessments of its veridicality, $F(1, 183) = 1.90$, $p = .17$. No other main or interaction effects were statistically significant, all Fs < 2.54, all ps > .11.

Given that the trivia questions had objectively correct answers, participants' prior knowledge might have affected conformity decisions. After all, had participants known the correct answer for certain, they would not have changed their answer, regardless of speech type or computer comments. A series of Chi-square analysis, however, showed that those who initially picked the correct answer were no less likely to switch to the computer's answer, suggesting that some participants picked the correct answer, not necessarily knowing that it was indeed the right one, all χ^2s < 1.97, all ps > .16.

Table 2
Means and standard deviations for evaluation of the computer.

	TTS	Recorded speech
Low rationals	4.84 (2.05)	5.92 (1.82)
High rationals	6.06 (2.43)	5.39 (1.81)
Low experientials	5.52 (2.55)	5.06(1.76)
High experientials	5.44 (2.13)	6.30 (1.68)

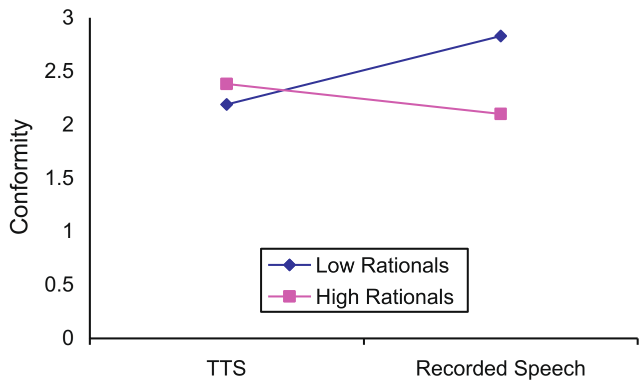

Fig. 2. Conformity by rationality and speech type.

6. Discussion

The present experiment examined (a) if social responses to computers are more likely to occur when the interface more closely resembles or somehow reminds of a human interactant and (b) if anthropomorphic interfaces elicit more positive reactions from those who are chronically less prone to engage in analytical thinking or rely more heavily on heuristic cues. Departing from previous HCI research (e.g., Cassell & Thorisson, 1999; Parise et al., 1999), which focused on if anthropomorphic interfaces make the interaction more pleasant and the system more trustworthy (anthropomorphism as an independent variable), the first question concerned if anthropomorphic interfaces facilitate the tendency to apply social expectations to computers (anthropomorphism as a moderator). Similarly, the second question pertained to when anthropomorphic interfaces would have greater impact, if any (interaction effect between anthropomorphism and users' cognitive style), rather than the main effect of the human likeness of the interface.

6.1. Theoretical implications

Overall, the current findings offered little support to the prediction that anthropomorphic interface will facilitate social responses to computers, and thus, amplify the previously found flattery effect. Quite contrarily, flattery significantly improved participants' overall perceptions of the computer only in the synthetic speech condition. Specifically, as compared to when factual feedback was provided, flattering comments garnered more positive social attractiveness and believability ratings, but such effects were confined to when a machine-like voice delivered the feedback. Possibly, just as the use of human face (vs. simple text) heightened social presence and induced a stronger social desirability bias in participants' responses (Sproull et al., 1996), recorded human speech might have reminded participants of a real person and triggered suspicion about his ulterior motivation behind flattery. Increased suspicion, in turn, might have induced negative feelings about the flatterer among some participants (Campbell & Kirmani, 2000), canceling out the otherwise positive effects of flattery. By contrast, when synthetic voice was used, participants might have become less concerned about the hidden motive of their interactant – after all, computers do not have intentions, good or bad – and thus, more positively inclined toward the computer that praised their performance. If so, one might suspect that the unanimously positive effects of flattery found in the previous study (Fogg & Nass, 1997) were because of, not despite, the fact that the flattery came from a computer, as opposed to a real person who might have self-centered motivation for ingratiation. Yet another possibility is that both speech conditions were considered to be sufficiently humanlike. That is, although those exposed to the recorded speech attributed significantly higher degrees of human likeness to the computer than those in the synthetic speech condition did, the difference in perceived anthropomorphism scores between the two was not large (4.99 vs. 5.78 on a 10-point scale). Perhaps, this relatively small difference in perceived human likeness might have suppressed the potential effects of anthropomorphic interface on the likelihood of social responses to computers, leading to the null effect.

Although the effects were not consistently observed for all dependent variables, users' cognitive style does seem to moderate the effects of anthropomorphic qualities embedded in the interface. Not only did low rationals evaluate the computer's performance more favorably when it used more humanlike than mechanical voices, but they were also more willing to switch to the computer's suggestions when the computer feedback was provided in recorded human speech. Likewise, high experientials not only rated the human-voiced computer's overall performance more positively than that of the synthetic-voiced one, but their self-confidence was dropped to the level of low experientials when the human-voiced computer invalidated their initial answer, suggesting that they took the computer's feedback more seriously. At the very least, these findings indicate that anthropomorphic interface has limited values in eliciting positive user reactions, operating only in conjunction with individuals' predispositions; that is, making the computer sound more humanlike, at least when not accompanied by significant functional improvement as in the current research, does not warrant more favorable user responses. As such, to better understand how anthropomorphic cues work in human–computer interaction, it seems imperative to incorporate various user characteristics in future research, over and above simple demographic factors or computing experience.

In addition, the present research identified when or for whom the surface-level anthropomorphism can be more effective: less analytical thinkers and more intuition-driven individuals. Given that the interaction content was held constant across conditions, the humanness of the computer voice should not have influenced their assessments of the computer's overall performance or conformity decisions, had such judgments been made on the strictly logical basis. And yet, the sound of voice significantly altered users' reactions when they are less prone to engage in effortful information processing or more willing to go by their instinct. Although this experiment focused on users' enduring dispositional qualities as a potential moderator of the effects of anthropomorphic interface, it seems worth examining if situational constraints, such as time pressure or multi-tasking, would similarly facilitate the effects of anthropomorphic interface as reported herein.

The interpretation that anthropomorphic interface was better received by low rationals and high experientials due to their susceptibility to peripheral cues comports well with other findings. That is, both low rationals and high experientials evaluated their own performance more positively when the computer threw in some meaningless, yet flattering, remarks, than when it offered factual feedback. Unlike the original flattery study (Fogg & Nass, 1997), in which no objective evaluation criteria were available for the participants' performance, the trivia game employed in this study had verifiable answers and the participants could tell whether they knew the answer or not with certain degrees of confidence. Even though the participants could have better gauged their performance, independent of the feedback they received from the computer, those chronically less motivated for effortful thinking and those relying heavily on their gut feelings thought they performed better when flattered.

It also merits note that users' affective and behavioral reactions do not operate in unison. For example, albeit only when carried by a synthetic voice, flattery enhanced participants' overall perceptions of social attractiveness and believability of the computer. However, flattery tended to *lower* participants' conformity to the computer's suggestions. Similarly, high experientials thought the human-sounding computer had offered more correct answers than the machine-like one did, and yet, they were no more likely to accept the computer's answers when interacting with the human-voiced computer. Taken together, these results suggest that affective reactions of users are not easily transferred to the behavioral domain, posing greater challenge to those purporting to induce behavioral changes by means of persuasive technology (Fogg, 2003).

6.2. Limitations and future directions

Some aspects of the present research might limit the generalizability of the findings. First, participants were explicitly told that the computer was programmed to generate random answers, for otherwise they would have invariably accepted the computer's suggestions. Such manipulation, however, could have made the participants process and respond to the computer messages somewhat differently than they would in other situations where the computer's correctness is reasonably assumed. Possibly, they might have become unusually suspicious, and thus, processed the computer messages more carefully than they would normally do. Alternatively, they might have paid little attention to the computer feedback, for it has no informational value. Although there seems to be no compelling reason to explain away the current findings on account of heightened suspicion or general dismissal of the computer output, future research should examine if the current results are replicable when people are not particularly encouraged to question the validity of information provided by the computer.

Second, this study investigated if anthropomorphic interfaces accentuate people's tendency to treat computers as social actors, using flattery effects as an example of social responses. Results showed that the use of human voice, as opposed to machine-generated one, failed to amplify flattery effects, but the way the anthropomorphic interface was operationalized might have contributed to such null findings. Specifically, the current research employed two different types of voices with varying degrees of human likeness, but the very presence of voice, regardless of its sound, might have been a cue powerful enough to trigger social responses to computers (Nass & Moon, 2000). Considering that speech is a uniquely human trait, the fact that the computer delivered its output in speech might have automatically activated social scripts associated with flattery, no matter how closely it approximates human voices. If so, one could argue that the current study does not constitute a fair test of the notion that anthropomorphism makes people more prone to project social expectations onto computers. To address this issue, future research should assess the validity of the 'anthropomorphism-as-amplifier-of-social responses' account by implementing different types of anthropomorphic attributes, such as visual images of static objects versus humans (Nowak & Rauh, 2005), which maximize the systematic variance in perceived human likeness.

Lastly, with respect to the manipulation of anthropomorphism, one might suspect that participants could have attributed different degrees of human likeness to the computer, depending on the comments it provided. That is, considering that ingratiation, as a means to acquire social approval or other kinds of rewards, requires social intelligence, the flattering computer might have been perceived as more humanlike than the one offering strictly factual, or computer-like, feedback. As such, it seems plausible to conceive of recorded versus synthetic speech variation as tapping *perceptual* anthropomorphism, while the flattery versus generic comments representing *behavioral* anthropomorphism (Nowak & Biocca, 2003). However, the current data showed that those having received flattering comments did not rate the computer as any more anthropomorphic than those presented with generic comments. In addition, if flattery indeed renders the computer more humanlike, or induces such impressions among users, it is all the more puzzling why then the match between perceptual and behavioral anthropomorphism did not amplify their effects; just as consistency between the human likeness of facial image and voice significantly improved users' evaluation of the system (Li & Nass, 2007), flattery effects should have been more pronounced when coupled with human than clearly machine-like voices. Nevertheless, the distinction and potential interplay between perceptual and behavioral, or formal and functional, anthropomorphism invites more systematic investigations in relation to how and why people make social attributions to the computer.

7. Conclusion

The present research found that although auditory resemblance to humans did not make people any more likely to evince social responses to flattering computers, it served as a peripheral cue that significantly improved users' evaluation of the computer's performance and conformity to the computer's suggestions among those chronically less inclined to think analytically. Similarly, those relying more heavily on their intuition showed greater sensitivity to the perceptual anthropomorphic cue embedded in the interface. Extending the previous research on human–computer interaction, which focused mostly on how various features of the computer interface, either source characteristics (e.g., computer gender) or message characteristics (e.g., flattery), alter user experience, the current study suggests the important role of receiver characteristics (e.g., cognitive style) in determining the nature of man–machine interaction. More than anything else, the present findings highlight the need for more nuanced and comprehensive approaches to this increasingly prevalent form of interaction, fully integrating various facets of communication processes.

Acknowledgement

This research was financially supported in part by the Institute of Communication Research, Seoul National University.

References

Bailenson, J. N., Blascovich, J., Beall, A. C., & Loomis, J. M. (2001). Equilibrium revisited: Mutual gaze and personal space in virtual environments. *Presence, 10*, 583–598.

Bailenson, J. N., Swinth, K., Hoyt, C., Persky, S., Dimov, A., & Blascovich, J. (2005). The independent and interactive effects of embodied-agent appearance and behavior on self-report, cognitive, and behavioral markers of copresence in immersive virtual environments. *Presence, 14*, 379–393.

Byrne, D. (1971). *The attraction paradigm*. New York: Academic Press.

Campbell, M. C., & Kirmani, A. (2000). Consumers' use of persuasion knowledge: The effects of accessibility and cognitive capacity on perceptions of an influence agent. *Journal of Consumer Research, 27*, 69–83.

Carli, L. L. (2001). Sex and social influence. *Journal of Social Issues, 57*, 725–741.

Cassell, J., & Thorisson, K. R. (1999). The power of a nod and a glance: Envelope vs. emotional feedback in animated conversational agents. *Applied Artificial Intelligence, 13*, 519–538.

Chen, S., & Chaiken, S. (1999). The heuristic-systematic model in its broader context. In S. Chaiken & Y. Trope (Eds.), *Dual-process theories in social psychology* (pp. 73–96). New York: The Guilford Press.

Dehn, D. M., & van Mulken, S. (2000). The impact of animated interface agents: a review of empirical research. *International Journal of Human–Computer Studies, 52*, 1–22.

Eagly, A. H., & Chaiken, S. (1975). An attribution analysis of the effect of communicator characteristics on opinion change: The case of communicator attractiveness. *Journal of Personality and Social Psychology, 32*, 136–144.

Epstein, S. (1994). An integration of the cognitive and the psychodynamic unconscious. *American Psychologist, 49*, 709–724.

Epstein, S., & Pacini, R. (1999). Some basic issues regarding dual-process theories from the perspective of cognitive-experiential self-theory. In S. Chaiken & Y. Trope (Eds.), *Dual-process theories in social psychology* (pp. 462–482). New York: The Guilford Press.

Fogg, B. J. (2003). *Persuasive technology: Using computers to change what we think and do*. San Francisco, CA: Morgan Kaufmann.
Fogg, B. J., & Nass, C. (1997). Silicon sycophants: the effects of computers that flatter. *International Journal of Human–Computer Studies, 46*, 551–561.
Heckman, C. E., & Wobbrock, J. O. (2000). Put your best face forward: Anthropomorphic agents, e-commerce consumers, and the law. *Agents 2000* (pp. 435–442). Barcelona, Spain.
Hinds, P. J., Roberts, T. L., & Jones, H. (2004). Whose job is it anyway? A study of human–robot interaction in a collaborative task. *Human–Computer Interaction, 19*, 151–181.
Lee, E.-J. (2003). Effects of "gender" of the computer on informational social influence: The moderating role of task type. *International Journal of Human–Computer Studies, 58*, 347–362.
Lee, E.-J. (2008a). Gender stereotyping of computers: Resource depletion or reduced attention? *Journal of Communication, 58*, 301–320.
Lee, E.-J. (2008b). Flattery may get computers somewhere, sometimes: The moderating role of output modality, computer gender, and user gender. *International Journal of Human–Computer Studies, 66*, 789–800.
Lee, K. M., & Nass, C. (2004). The multiple source effect and synthesized speech: Doubly disembodied language as a conceptual framework. *Human Communication Research, 30*, 182–207.
Lee, E.-J., Nass, C., & Brave, S. (2000). Can computer-generated speech have gender? An experimental test of gender stereotype. In *CHI extended abstracts* (pp. 289–290). New York: ACM Press.
Li, G., & Nass, C. (2007). When a talking-face computer agent is half-human and half-humanoid: Human identity and consistency preference. *Human Communication Research, 33*, 163–193.
MacDorman, K. F. (2006). Subjective ratings of robot video clips for human likeness, familiarity, and eeriness: An exploration of the uncanny valley. In *ICCS/CogSci-2006 long symposium: Toward social mechanisms of android science*. July 26, 2006. Canada: Vancouver.
Moon, Y. (2000). Intimate exchanges: Using computers to elicit self-disclosure from consumers. *Journal of Consumer Research, 26*, 323–339.
Moon, Y., & Nass, C. (1996). How "real" are computer personalities?: Psychological responses to personality types in human-computer interaction. *Communication Research, 23*, 651–674.
Mori, M. (1970). Bukimi no tani [the uncanny valley]. *Energy, 7*, 33–35.
Nass, C., & Brave, S. (2005). *Wired for speech: How voice activates and advances the human-computer relationship*. Cambridge, MA: MIT Press.
Nass, C., & Moon, Y. (2000). Machines and mindlessness: Social responses to computers. *Journal of Social Issues, 56*, 81–103.
Nowak, K. L., & Biocca, F. (2003). The effect of the agency and anthropomorphism on users' sense of telepresence, copresence, and social presence in virtual environments. *Presence: Teleoperators and Virtual Environments, 12*, 481–494.
Nowak, K. L., & Rauh, C. (2005). The influence of the avatar on online perceptions of anthropomorphism, androgyny, credibility, homophily, and attraction. *Journal of Computer-Mediated Communication*, *11*(1), article 8. Retrieved July 20, 2009. Available from http://jcmc.indiana.edu/vol11/issue1/nowak.html.
Pacini, R., & Epstein, S. (1999). The relation of rational and experiential information processing styles to personality, basic beliefs, and the ratio-bias phenomenon. *Journal of Personality and Social Psychology, 76*, 972–987.
Parise, S., Kiesler, S., Sproull, L., & Waters, K. (1999). Cooperating with life-like interface agents. *Computers in Human Behavior, 15*, 123–142.
Petty, R. E., & Cacioppo, J. T. (1986). *Communication and persuasion: Central and peripheral routes to persuasion*. New York: Springer-Verlag.
Powers, A., Kiesler, S., 2006. The advisor robot: tracing people's mental model from a robot's physical attributes. In *Proceedings of the 1st ACM SIGCHI/SIGART conference on human–robot interaction* (pp. 218–225). Salt Lake City, Utah, USA.
Reeves, B., & Nass, C. (1996). *The media equation: How people treat computers, television, and new media like real people and places*. New York: Cambridge University Press.
Sproull, L., Subramani, M., Kiesler, S., & Walker, J. H. (1996). When the interface is a face. *Human–Computer Interaction, 11*, 97–124.
Sundar, S. S. (2004). Loyalty to computer terminals: Is it anthropomorphism or consistency? *Behaviour & Information Technology, 23*, 107–118.
Yee, N., Bailenson, J. N., & Rickertsen, K. (2007). A meta-analysis of the impact of the inclusion and realism of human-like faces on user experiences in interfaces. In *Proceedings of the conference on human factors in computing systems (ACM SIGCHI)* (pp. 1–10). New York: ACM Press.