Computers in Human Behavior 97 (2019) 304–316

Contents lists available at ScienceDirect

Computers in Human Behavior

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

Full length article

Humanizing chatbots: The effects of visual, identity and conversational cues on humanness perceptions

Eun Go[a,*], S. Shyam Sundar[b]

[a] Department of Broadcasting & Journalism, College of Fine Arts & Communication, Western Illinois University, 318 Sallee Hall, 1 University Cir, Macomb, IL 61455-1390, USA

[b] Media Effects Research Laboratory, Donald P. Bellisario College of Communications, The Pennsylvania State University, 122 Carnegie Building, University Park, PA 16802-5101, USA

ARTICLE INFO

Keywords:
Online chat agents
Message interactivity
Identity cue
Anthropomorphic visual cue
Compensation effect
Expectancy violation effect

ABSTRACT

Chatbots are replacing human agents in a number of domains, from online tutoring to customer-service to even cognitive therapy. But, they are often machine-like in their interactions. What can we do to humanize chatbots? Should they necessarily be driven by human operators for them to be considered human? Or, will an anthropomorphic visual cue on the interface and/or a high-level of contingent message exchanges provide humanness to automated chatbots? We explored these questions with a 2 (anthropomorphic visual cues: high vs. low anthropomorphism) × 2 (message interactivity: high vs. low message interactivity) × 2 (identity cue: chat-bot vs. human) between-subjects experiment ($N = 141$) in which participants interacted with a chat agent on an e-commerce site about choosing a digital camera to purchase. Our findings show that a high level of message interactivity compensates for the impersonal nature of a chatbot that is low on anthropomorphic visual cues. Moreover, identifying the agent as human raises user expectations for interactivity. Theoretical as well as practical implications of these findings are discussed.

Interactive chat agents have now become quite common, with several sites and applications using a variety of chatbots, ranging from customer-service agents on e-commerce sites to Apple's Siri and Amazon's Echo. The main functions of online chat agents are to interact with users, respond to their questions, and address their concerns. The experience provided by these agents is considered to be better than static delivery of information, such as a list of frequently asked questions (FAQs), because the agents offer more interactive delivery of messages to users, responding specifically to their questions – so much so that several companies have begun to use these agents as replacements for human agents or telephone-based call support. As such, the use of online chat agents seems to hold tremendous promise for providing users with quick and convenient support. However, it has also resulted in new challenges. In particular, online chat agents are short on humanness, considering that their conversation styles tend to be artificial and somewhat stilted because of their adherence to programmed scripts.

Nonetheless, if online chat agents assume roles hitherto fulfilled by humans, it is necessary to make their interactions as similar as possible to those of human beings. There are generally three ways to suggest humanness among online chat agents. The first could be the use of human figures (visual cues). Second, human-associated names or identity (identity cues) can be used for chat agents. The last way is mimicking of human languages (conversational cues). Therefore, this study tested the distinct and combined effects of three types of cues that potentially enhance humanness of chat agents. The findings will not only advance our theoretical knowledge about the psychological effects of different human-like characteristics of interactive online chat agents, but will also provide practical guidelines on how such characteristics should be implemented to online chat agents in order to create the most desirable outcomes for users in a number of domains, from customer service in e-commerce to counseling services in e-health.

1. What makes a chat agent human-like?

During the last half-century, scholars and designers have dedicated considerable effort to make intelligent chat agents more similar to humans. The easiest way to enhance the humanness of chat agents might be the use of human figures. Human visual (anthropomorphic) cues of agents can be richly suggestive and can shape social perceptions, due to mindlessness, priming and over-attribution (Kim, 2010; Lee & Oh, 2015). Human-like visual cues are likely to trigger

* Corresponding author.
E-mail addresses: e-go@wiu.edu (E. Go), sss12@psu.edu (S.S. Sundar).

https://doi.org/10.1016/j.chb.2019.01.020
Received 28 June 2018; Received in revised form 6 January 2019; Accepted 19 January 2019
Available online 29 January 2019

"humanness" heuristics (Sundar, 2008), leading users to treat chat agents, such as chatbots, as human and act socially towards them (Baylor & Kim, 2003; Gong & Nass, 2007; Kim & Sundar, 2012; Nowak, 2004). Therefore, simple interface-level manipulation of chat agents with visual cues can suggest humanness to users.

Next, the use of human names or identities is another easy way to enhance the humanness of chat agents. This would be an effective enhancement because of our tendency to perceive things by their labels. Cognitive psychologists have emphasized the importance of category-based perceptions activated by the social labels (or identity cues) assigned to individuals or objects, arguing that individuals tend to use major attributes attached to labels in order to minimize cognitive effort when making judgments or in forming impressions of others (Ashforth & Humphrey, 1997; Gelman & Heyman, 1999). Such labels, then, evoke certain stereotypes or perceptions from one's schema, allowing one to make heuristic judgments based on available cues. Tversky and Kahneman (1974) also explained the use of heuristic in judgments in that individuals tend to make estimates by starting form an initial value. For example, labels (identity cues) attached to chat agents can serve as an initial value and then other characteristics can be considered in making judgment on the humanness of chat agents. Offering the user information about whether the agent is a real human or a chatbot may elicit machine vs. human heuristics that can be related to anthropomorphic perceptions, according to the Modality-Agency-Interactivity-Navigability (MAIN) model, which posits that interface cues shape user perceptions by triggering cognitive heuristics about the nature and substance of the interaction (Sundar, 2008). If an agent is presumed to be a chatbot, then users are more likely to evaluate the quality of the agent's performance based on their preexisting perceptions of chatbots or machines. Human identity cues, on the other hand, will predispose users to evaluate the quality of the agent's performance based on their expectations of humans. That is, identity cues suggesting that the user is chatting with a human agent or machine agent can trigger human or machine heuristics respectively and accordingly affect the criteria by which they evaluate the quality of the interaction. As Koh and Sundar (2010) have argued, the actual conversational quality of the online chat agent might not matter in evaluating the agent's performance if the identity assigned to the agent already elicits stereotypical judgments.

Aside from identifying the chat agent as human or providing humanlike cues on the interface, another way to enhance chatbot humanness would be the use of human language. This option, however, presents more challenges than just using human figures or human names (identity) due to the complexities of human language. Scientists therefore have invested much effort into designing discourse that have human-like characteristics (Ghose & Barua, 2013). A key characteristic of human communication is "contingency" in responses. In other words, responses exchanged in a human dialogue are interconnected with each other (Sundar, Bellur, Oh, Jia, & Kim, 2016). This means that a response is contingent upon the preceding message as well those preceding it, in a threaded conversations (Rafaeli, 1988). In the literature on human-computer interaction, such contingency is referred to as "message interactivity" (Sundar, 2009). Therefore, message interactivity in a chat context can be conceptualized as the level of contingency (or interdependent threadedness) in message exchanges. For example, when two individuals are having a conversation, if one not only acknowledges the other's message but also shows an awareness of previous conversations when responding to the message in a back-and-forth fashion, then the conversation follows an extended thread and is considered fully interactive or responsive (Rafaeli, 1988). Such message-interactivity would make an online chat agent appear "human-like" because it mimics the contingency found in message exchanges between human interactants (Rafaeli, 1988).

2. Perceptual effects of the three factors of humanness

Whether all three factors work well for suggesting humanness or not can be dictated by diverse psychological outcomes. First, it has been found that anthropomorphic visual presentation increases the sense of social presence among users who are having conversations with such agents. Social presence is formally defined as the "degree of salience of the other person in the interaction" (Short, Williams, & Christie, 1976, p. 65). It is also defined as the feeling of "being with another in a mediated environment" (Biocca, Harms, & Burgoon, 2003, p. 14). Highly anthropomorphic visual cues (e.g., a human figure) could increase the awareness or salience of the "other person" than low anthropomorphic visual cues (e.g., a bubble) because the presence of a human figure attached to a chat agent itself can be suggestive of the existence of the "other person" in the interaction. Indeed, Kim and Sundar (2012) found that participants in human-like agent conditions show greater feelings of interacting with others. Cyr, Head, Larios, and Pan (2009) also identified that adding human images convey the feelings of human contact. The anthropomorphic visual cues of an online chat agent is also expected to increase perceived homophily, which is defined as "the amount of similarity two people perceive themselves as having" (Rocca & McCroskey, 1999, p. 309), since human-like figure of the agent is likely to be perceived as being more similar to the user than a bubble figure.

A human identity cue is also expected to increase social presence and perceived homophily. As mentioned earlier, offering the user information about whether the agent is a real human or a chat-bot may elicit machine vs. human heuristics that can be related to anthropomorphic perceptions. As Koh and Sundar (2010) have argued, if an agent is presumed to be operated by a chat-bot, then users are more likely to evaluate the quality of the agent's performance based on their pre-existing perceptions of chat-bots or machines elicited by that identity cue, regardless of the agent's actual performance quality. Human identity cues as opposed to chat-bot identity cues are expected to make users evaluate the quality of the agent's performance based on their expectations of the characteristics of human agents. Sundar et al. (2016) have shown that participants in a human agent condition evaluate the chat agent as more human-like, natural, and life-like than those in a chat-bot agent condition. If online users perceive online chat agents with human identity cues as human-like, then participants in a human identity cue condition will show a greater sense of social presence than those in a chat-bot identity cue condition because human-like perception can make users perceive a chat agent as "another person." By the same token, human-like perception formed by a human identity cue makes users being more likely to perceive similarity to the agent that has a human identity cue as apposed to a chat-bot identity cue.

Last, a conversational cue (message interactivity) of online chat agents can also facilitate a feeling of interacting with other people (i.e., social presence) without physical co-presence. In other words, by engaging in back-and-forth message exchanges, two interactants are able to recognize one another even in the online chatting context. Sundar, Go, Kim, and Zhang (2015) found that higher message interactivity in a chatting context heightens a feeling of the other's presence. Furthermore, due to the back-and-forth nature of high message interactivity, as with contingency, people are more likely to perceive it as a dialogue, which is a core characteristic of human-to-human communication. Therefore, it would allow an online chat agent to appear to have a "human-like voice." Therefore, a chat agent that carries over a high level of contingency in the exchange of messages is more likely to be perceived as similar to human beings which will, in turn, increase homophily perception.

Conversational cues (message interactivity) of online chat agents are expected to lead additional unique psychological outcomes. As mentioned earlier, a key element of message interactivity is its ability to deliver contingent messages. Thus, given the back-and-forth nature of

high message interactivity, responsive conversations, operated by high message interactivity, are more likely to deliver high levels of contingency than noninteractive or reactive conversations (Hrastinski, 2008; Sundar et al., 2010, 2015; 2016). Like perceived contingency, when the conversation is considered fully interactive or responsive (Rafaeli, 1988), individuals will perceive greater levels of dialogue emerging from threaded message exchanges with a chat agent during computer-mediated communication (Bellur & Sundar, 2017).

The aforementioned psychological outcomes are significant in that they are critical mediators of attitudinal and behavioral outcomes. First, social presence has been demonstrated as a key predictor of favorable attitudes and behaviors, because social presence has psychological elements of emotional closeness and/or social connectedness (Bente, Rüggenberg, Krämer, & Eschenburg, 2008; Biocca et al., 2003; Palmer, 1995; Rice, 1984). Therefore, the greater a user experience social presence while interacting with an online chat agent, the greater a user feels emotional closeness and/or social connectedness. Thus, such heightened feeling of connectedness can result in positive evaluations of the agent. The positive attitudes toward the agents subsequently engender favorable attitudes toward the website where the agent is embedded, and, ultimately, creates behavioral intentions among users to visit the websites again in the future.

Next, individuals have a propensity to evaluate those similar to themselves favorably (Goethals & Nelson, 1973). The more similar two people are to each other, the more they attempt to communicate with and understand each other (Rogers & Bhowmik, 1970). That's why perceived homophily comes into play in creating favorable attitudes and behavioral intentions. Therefore, if users perceived an online chat agent as being more similar to themselves, they are likely to evaluate the agent (or the website where the agent is embedded) more favorably. Indeed, studies have found that an online chat agent with human-like morphology is likely to be rated as more likable than an agent with less human-like morphology (Koda, 1996; Wexelblat, 1998). Favorable attitudes are also likely to result in greater behavioral intentions to revisit the interaction.

Last, Sundar et al. (2016) have documented that the message interactivity occurring through live-chatting fosters positive attitudes and increases behavioral intentions to return to a given website, by way of imbuing users with perceptions of contingency. Bellur and Sundar (2017) have also shown the mediating effects of perceived contingency between message interactivity and users' positive attitudes toward messaging systems. In a similar vein, perceived dialogue cued by higher levels of contingency on interfaces may also lead to positive attitudinal and behavioral outcomes. Dialogue perception is critical in two-way communication (Kent & Taylor, 2002; Seltzer & Mitrook, 2007), as it may make online interactions feel like face-to-face conversations, in turn creating positive attitudes toward the agent and website that offer dialogic communication and behavioral intentions to revisit the site.

In sum, this study suggests the following hypotheses in regard to the anthropomorphic visual cue, identity cue, and conversational style of an online chat agent on psychological, attitudinal, and behavioral outcomes.

3. Proposed hypotheses

H1. Higher levels of anthropomorphic visual cues in an online chat agent will lead to (a) greater feelings of social presence and (b) greater perceived homophily.

H2. Greater feelings of social presence achieved through the anthropomorphic visual cues of an online chat agent will lead to (a) more favorable attitudes, and (b) greater behavioral intentions to return to a given website.

H3. Greater perceived homophily achieved through the anthropomorphic visual cues of an online chat agent will lead to (a) more favorable attitudes, (b) greater behavioral intentions to return to a given website.

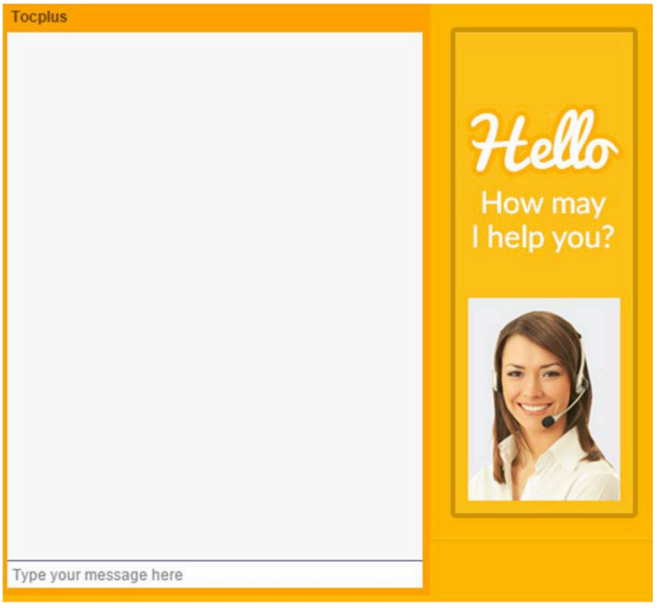

Fig. 1. A live chat window with a highly anthropomorphic visualized online chat agent.

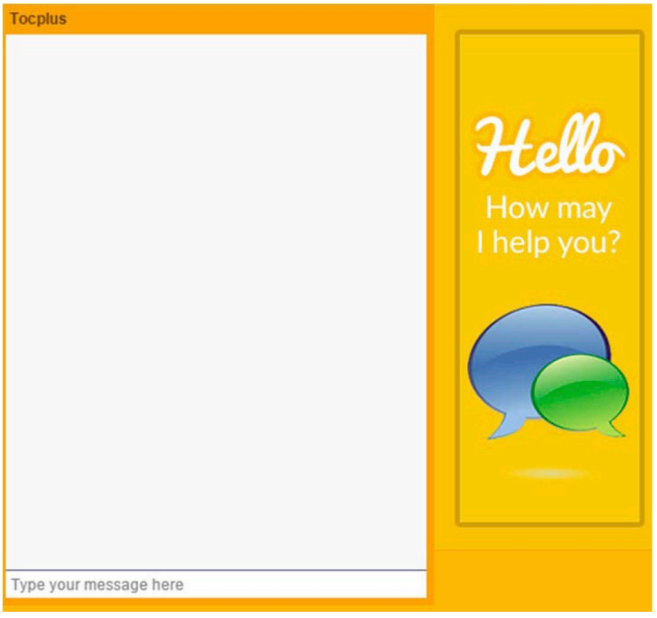

Fig. 2. A live chat window with a low anthropomorphic visualized online chat agent.

Table 1
Mean and standard deviation of social presence and perceived homophily.

Outcome variables	Anthropomorphic Visual Cue			
	High		Low	
	M	SD	M	SD
Social presence	5.02	0.18	4.95	0.18
Perceived homophily	4.39	0.16	4.34	0.16

H4. Human identity cues of online chat agents will lead to greater feelings of social presence and (b) greater perceived homophily.

Table 2
Main effects of anthropomorphic visual cue on outcome variables.

Outcome variables	F-value	p-value	Partial Eta Squared
Social presence	0.07	$p = .80$	0.00
Perceived homophily	0.04	$p = .83$	0.00

H5. Greater feelings of social presence achieved through the human identity cues of online chat agents will lead to (a) more favorable attitudes, and (b) greater behavioral intentions to return to a given website.

H6. Greater perceived homophily achieved through the human identity cues of online chat agents will lead to (a) more favorable attitudes, and (b) greater behavioral intentions to return to a given website.

H7. Higher message interactivity will lead to (a) greater feelings of social presence, (b) greater perceived homophily, (c) greater perceived contingency, and (d) greater perceived dialogue.

H8. Greater feelings of social presence achieved through higher message interactivity will lead to (a) more favorable attitudes, and (b) greater behavioral intentions to return to a given website.

H9. Greater perceived homophily achieved through higher message interactivity will lead to (a) more favorable attitudes, and (b) greater behavioral intentions to return to a given website.

H10. Higher levels of perceived contingency achieved through higher message interactivity will lead to (a) more favorable attitudes and (b) greater behavioral intentions to return to a given website.

H11. Higher levels of perceived dialogue achieved through higher message interactivity will lead to (a) more favorable attitudes and (b) greater behavioral intentions to return to a given website.

It must be noted, however, that it is uncertain how well all three humanness cues work together to optimize user experience with a chat agent. Each has been demonstrated as working well to improve humanness of chatbots. However, it is unknown that how they interact with each other when they are implemented together in a chatbot. Therefore, we propose the following research question for study.

RQ1: Are there interaction effects among the three humanness factors on psychological outcomes related to user experience of an online chat agent?

4. Method

This study employed a 2 (Anthropomorphic visual cue: low anthropomorphic vs. high anthropomorphic) × 2 (Identity cue: chat-bot vs. human) × 2 (Message Interactivity: low vs. high) between-participants factorial experiment.

4.1. Participants

All study participants (N = 141) were recruited via Amazon's Mechanical Turk, "a marketplace for work" where virtual workers perform "human intelligence tasks" (HITs). Previous studies have shown that the samples recruited from Amazon's Mechanical Turk (AMT) tend to be more demographically heterogeneous than typical American college samples (e.g., Buhrmester, Kwang, & Gosling, 2011; Goodman, Cryder, & Cheema, 2013). As Mason and Suri (2012) pointed out, workers in Mechanical Turk tend to be "from a very diverse background, spanning a wide range of age, ethnicity, socio-economic status, language, and country of origin" (p. 2). The use of AMT has been validated as a tool for conducting research using diverse methods, from surveys (Buhrmester et al., 2011; Gosling, Vazire, Srivastava, & John, 2004) to behavioral experiments (Mason & Watts, 2009; Suri & Watts, 2011). To ensure quality responses, participants needed to have HIT approval ratings of above 90% in order to be eligible for our study. All were United States residents and paid $1 each for participation. The demographics of the recruited participants were as follows: the majority of the participants were Caucasian (56%), the average age was 33.91 (SD = 10.09), and the ratio of males and females was almost equal (male: 50.4%, female: 49.6%). In terms of education, most of the participants (85.1%) had completed college. Participants were randomly assigned to one of eight versions of the same website, varied systematically to investigate the three independent variables.

4.2. Procedure

At the time of recruitment, participants signed up for a specific timeslot, much like in a lab experiment. Once they accessed the study site at the time, they were provided an informed consent form to approve before moving forward with their participation in the study. After they consented, they were led to an online pre-questionnaire to measure their demographic information. Next, they were asked to read a scenario–each participant was told that he or she was choosing the best digital camera to purchase as a birthday gift for a friend based on the friend's preferences, which were spelled out (brand, budget, camera type and size).

After giving the participants time to read their scenarios (see Appendix A), participants were asked to explore "Digital World," a mock website that ostensibly sells digital cameras. They were then asked to interact with the online chat agent using the website's chat function by clicking the "Live Chat" icon displayed in the top-right corner of the website. Once a participant initiated a chat, the system said "A sales associate [a chat-bot] will be with you shortly." Then, the agent introduced itself and proceeded to ask questions. Given that the online chat agent needed to know what the customer needed, the chat followed the format of a Q&A. The questions centered on the purpose of buying a digital camera, as well as the customer's preferences regarding camera type, brand, and budget. Once a question was asked, the agent waited until the participant answered. Then, the agent responded to the participant's answer. If participants did not follow the scenario strictly when answering questions, or did not focus on the chatting, their answers were excluded from data analysis. After asking four questions, the agent made a recommendation based on the information provided. By instructing participants to use the information provided in their scenarios when answering the agent's questions, the study was able to ensure that the content of the chats was controlled. At the end of the chat, the system provided a survey link to a post-experiment questionnaire.

Table 3
Significant indirect effects of visual cues on attitudinal and behavioral outcome via social presence.

Mediation Path	Indirect Effect Bootstrap Estimate (*b*)	Indirect Effect 95% Confidence Intervals	
		LL	UL
Visual cue → Social presence → Perceived expertise	−0.01	−0.13	0.08
Visual cue → Social presence → Perceived friendliness	−0.04	−0.34	0.24
Visual cue → Social presence → Website attitudes	−0.02	−0.19	0.11
Visual cue → Social presence → Behavioral intention	−0.03	−0.28	0.17

Table 4
Significant indirect effects of identity cues on attitudinal and behavioral outcome via perceived homophily.

Mediation Path	Indirect Effect Bootstrap Estimate (*b*)	Indirect Effect 95% Confidence Intervals	
		LL	UL
Visual cue → Homophily → Perceived expertise	0.01	−0.09	0.16
Visual cue → Homophily → Perceived friendliness	0.02	−0.12	0.19
Visual cue → Homophily → Website attitudes	0.02	−0.12	0.20
Visual cue → Homophily → Behavioral intention	0.02	−0.17	0.25

Table 5
Mean and standard deviation of social presence and perceived homophily.

Outcome variables	Identity Cue			
	High		Low	
	M	SD	M	SD
Social presence	5.16	0.18	4.81	0.18
Perceived homophily	4.38	0.16	4.36	0.16

Table 6
Main effects of identity cue on outcome variables.

Outcome variables	F-value	p-value	Partial Eta Squared
Social presence	1.85	$p = .18$	0.01
Perceived homophily	0.00	$p = .95$	0.00

4.3. Stimulus material

A prototype of a commercial website was constructed for the experiment. A digital camera was chosen as the target product because it has relatively complex product features and a large number of choices; therefore, individuals are likely to need help from an online chat agent in selecting the appropriate product. On the first page of the website, a number of digital cameras were presented. Specifically, digital cameras were displayed in three categories: DSLR cameras, compact digital cameras, and point-and-shoot digital cameras. There was also a live-chat button at the top right corner of the first page. If a participant clicked on the live-chat button, s/he saw the chat window pop up. The online chat agent then appeared in the chat window. Participants' answers as well as responses from the system were displayed in the chat box.

4.4. Manipulation of visual cues

This study followed the approach of manipulating anthropomorphic visual cues used in previous studies (e.g., Cyr et al., 2009; Kim & Sundar, 2012). For those participants in the high anthropomorphic visual cue condition, the online chat agent that appeared in the chatting window was a picture of an actual person (Fig. 1), while for those participants in the low anthropomorphic visual cue condition, the online chat agent appeared as a dialog bubble figure (Fig. 2). Participants were continuously exposed to the manipulation of the online chat agent's visual cue while chatting.

4.5. Manipulation of the identity cue

In order to manipulate the identity cue, participants were informed whether the online chat agent was operated by a human or a chat-bot. This method of manipulating identity cues has been used in previous studies (e.g., Koh & Sundar, 2010; Sundar et al., 2016). In the human identity condition, the system announced, “A sales associate will be with you shortly.” Then the online chat agent initiated the conversation by saying, “Hi! I'm Alex, a sales associate.” On the other hand, in the chat-bot identity condition, the system announced, “An automated chat-bot will be with you shortly.” Then the agent introduced itself by saying, “Hi! I'm Alex, an automated chat-bot.”

4.6. Manipulation of message interactivity

Our approach of manipulating message interactivity was based on methods used in previous studies (e.g., Bellur & Sundar, 2017; Sundar et al., 2016). The experimental manipulation of two levels of message interactivity was performed using the chatting function offered by the site. Each participant in the low message-interactivity condition engaged in a simple back-and-forth exchange with the agent. Specifically, the online chat agent did not acknowledge the participant's responses. In this case, then, there was no threadedness in the exchange between the agent and participant. For example, if the online chat agent asked: “Is this purchase for you or is this a gift?” and the participant answered: “This is for a gift,” the agent responded with: “Okay.”

On the other hand, in the high message-interactivity condition, each participant and the online chat agent engaged in a more contingent message exchange. The online chat agent was more responsive to the participant's messages by acknowledging the participant's responses to previous questions. For example, the online chat agent first asked whether this purchase was for the customer or a gift. The participant answered that the purchase was a gift. In the next question, the agent asked what type of camera the customer hoped to purchase. The participant answered that he or she was looking for a point-and-shoot camera. Then the agent responded: “Okay, so you are looking for a point-and-shoot camera as a gift. Let me ask another question.” As the above example suggests, in the high message-interactivity condition, the online chat agent's responses not only acknowledged the preceding question, but also showed an awareness of the participant's previous responses. Participants in both conditions chatted with the same human confederate as done in other similar studies (e.g., Sundar et al., 2016) to

Table 7
Significant indirect effects of identity cues on attitudinal and behavioral outcome via social presence.

Mediation Path	Indirect Effect Bootstrap Estimate (*b*)	Indirect Effect 95% Confidence Intervals	
		LL	UL
Identity cue → Social presence → Perceived expertise	0.05	−0.02	0.17
Identity cue → Social presence → Perceived friendliness	0.19	−0.10	0.52
Identity cue → Social presence → Website attitudes	0.10	−0.03	0.31
Identity cue → Social presence → Behavioral intention	0.14	−0.06	0.40

Table 8
Significant indirect effects of identity cues on attitudinal and behavioral outcome via perceived homophily.

Mediation Path	Indirect Effect Bootstrap Estimate (*b*)	Indirect Effect 95% Confidence Intervals	
		LL	UL
Identity cue → Homophily → Perceived expertise	−0.00	−0.13	0.10
Identity cue → Homophily → Perceived friendliness	−0.00	−0.15	0.14
Identity cue → Homophily → Website attitudes	−0.00	−0.17	0.14
Identity cue → Homophily → Behavioral intention	−0.00	−0.21	0.19

Table 9
Mean and standard deviation of social presence, perceived homophily, contingency, and dialogue.

Outcome variables	Message Interactivity			
	High		Low	
	M	SD	M	SD
Social presence	5.25	0.18	4.72	0.18
Perceived homophily	4.38	0.17	4.36	0.17
Perceived contingency	5.50	0.16	4.62	0.16
Perceived dialogue	5.95	0.13	5.12	0.13

Table 10
Main effects of identity cue on outcome variables.

Outcome variables	F-value	p-value	Partial Eta Squared
Social presence	4.24	$p < .05$	0.03
Perceived homophily	4.18	$p < .05$	0.03
Perceived contingency	14.29	$P < .001$	0.10
Perceived dialogue	19.37	$P < .001$	0.13

manipulate message interactivity. In addition, by allowing only five M-Turk workers to access each experimental session, the confederate was able to chat with all the participants individually.

4.7. Measures

All of the following measures, except the manipulation check measure for the identity cue, were assessed on a 7-point Likert scale.

Manipulation checks. For the visual cue manipulation check, the perceived anthropomorphism of the online chat agent was assessed with the following items: the agent's profile picture "does not look human/looks very human," "does not look realistic/looks very realistic," and "looks very cartoon-like/does not look like a cartoon" (Nowak & Rauh, 2005) ($M = 4.37$, $SD = 1.92$, Cronbach's $\alpha = 0.84$). For the identity manipulation check, participants were asked whether they had chatted with either "an automated chat-bot" or "a sales associate." In order to check for the message interactivity manipulation, four items were used. The items were as follows: "The chat agent remembered my responses," "The chat agent's responses were related to my earlier responses," "The agent took into account my previous interactions with it," and "The chat agent gave some smart suggestions based on my responses." These items were based on those used by Bellur and Sundar (2017) ($M = 5.53$, $SD = 1.47$, Cronbach's $\alpha = 0.94$).

Outcome variables. Social presence was measured with five items adopted from a study by Gefen and Straub (2003) that measured psychological sense of others without co-presence ($M = 5.00$, $SD = 1.52$, Cronbach's $\alpha = 0.96$). Ratings of perceived homophily were measured with four items adopted from Nowak and Rauh (2005). The items ask whether the online chat agent is very similar to me, thinks a lot like me, is very like me and behaves a lot like me ($M = 4.37$, $SD = 1.47$, Cronbach's $\alpha = 0.92$). Perceived contingency was assessed via five items adopted from Sundar et al. (2015), including whether participants perceived that the chat agent carefully registered their responses and gave feedback and whether the interaction with the chat agent was like a continuous thread ($M = 5.05$, $SD = 1.52$, Cronbach's $\alpha = 0.90$). To measure perceived dialogue, five items, for example, "I felt the chat agent considered its customers' unique requests," and "I felt like the chat agent was engaged in an active dialogue with their customers," were used (Sundar et al., 2016) (M = 5.53, SD = 1.27, Cronbach's $\alpha = 0.92$).

As measured in Koh and Sundar (2010), participants' evaluations on the chat agent were measured according to two factors: perceived expertise and perceived friendliness. The extent to which the participants perceived the agent as intelligent, knowledgeable, competent, and informed was indexed to determine perceived expertise ($M = 5.87$, $SD = 1.16$, Cronbach's $\alpha = 0.92$). The extent to which the participants perceived the agent as empathetic, personal, warm, emotionally invested, willing to listen, careful, and open was indexed to determine perceived friendliness ($M = 4.72$, SD = 1.50, Cronbach's $\alpha = 0.94$). Attitudes toward the website were measured on a scale (anchored between "describes very poorly" and "describes very well") featuring 10 adjectives (appealing, useful, positive, good, attractive, exciting, pleasant, likeable, interesting and high-quality), adopted from Sundar, Oeldorf-Hirsch, Nussbaum, and Behr (2011) ($M = 5.33$, $SD = 1.60$, Cronbach's $\alpha = 0.96$). To measure individuals' behavioral intentions to return to the site, six Likert-type items, for example, "I would bookmark this website for future use," "I would visit this website again in the future," obtained from Hu and Sundar (2010), were used ($M = 4.95$, $SD = 1.60$, Cronbach's $\alpha = 0.95$).

5. Results

5.1. Manipulation check

Anthropomorphism manipulation check. The results of an independent sample *t*-test confirmed that participants in the high

Table 11
Significant indirect effects of message interactivity on attitudinal and behavioral outcome via social presence.

Mediation Path	Indirect Effect Bootstrap Estimate (*b*)	Indirect Effect 95% Confidence Intervals	
		LL	UL
Message cue → Social presence → Perceived expertise	0.04	0.00	0.13
Message cue → Social presence → Perceived friendliness	0.27	0.03	0.56
Message cue → Social presence → Website attitudes	0.10	0.02	0.25
Message cue → Social presence → Behavioral intention	0.18	0.03	0.42

Table 12
Significant indirect effects of message interactivity on attitudinal and behavioral outcome via perceived homophily.

Mediation Path	Indirect Effect Bootstrap Estimate (*b*)	Indirect Effect 95% Confidence Intervals	
		LL	UL
Message cue → Homophily → Perceived expertise	0.06	0.00	0.17
Message cue → Homophily → Perceived friendliness	0.10	0.00	0.28
Message cue → Homophily → Website attitudes	0.08	0.00	0.24
Message cue → Homophily → Behavioral intention	0.15	0.01	0.39

Table 13
Significant indirect effects of message interactivity on attitudinal and behavioral outcome via perceived contingency.

Mediation Path	Indirect Effect Bootstrap Estimate (*b*)	Indirect Effect 95% Confidence Intervals	
		LL	UL
Message cue → Contingency → Perceived expertise	−0.09	−0.27	0.02
Message cue → Contingency → Perceived friendliness	−0.03	−0.19	0.13
Message cue → Contingency → Website attitudes	−0.08	−0.31	0.06
Message cue → Contingency → Behavioral intention	−0.08	−0.37	0.13

Table 14
Significant indirect effects of message interactivity on attitudinal and behavioral outcome via perceived dialogue.

Mediation Path	Indirect Effect Bootstrap Estimate (*b*)	Indirect Effect 95% Confidence Intervals	
		LL	UL
Message cue → Dialogue → Perceived expertise	0.46	0.23	0.78
Message cue → Dialogue → Perceived friendliness	0.30	0.12	0.61
Message cue → Dialogue → Website attitudes	0.59	0.30	0.79
Message cue → Dialogue → Behavioral intention	0.44	0.18	0.89

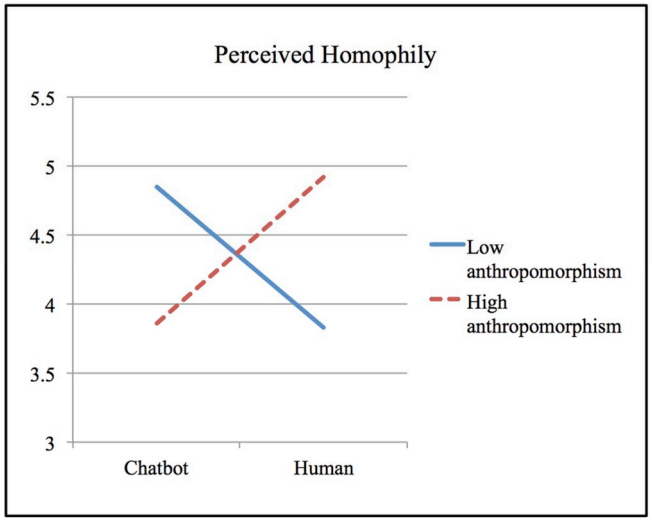

Fig. 3. Interaction effect of visual cue and identity cue on perceived homophily.

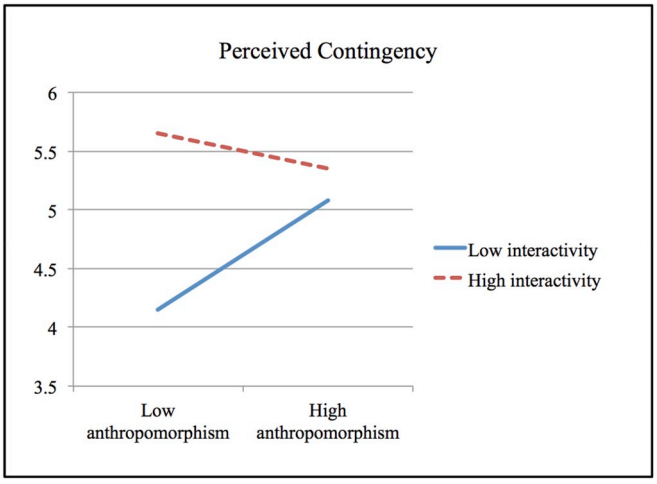

Fig. 4. Interaction effect of visual cue and message interactivity on perceived contingency.

anthropomorphic condition perceived the agent to be more realistic and human-like, and not cartoon-like (M = 5.78, SD = 1.48), than those participants in the low anthropomorphic condition (M = 3.00, SD = 1.17), $t(139) = -12.32$, $p < .001$, Cohen's d = 2.08).

Identity manipulation check. In order to check for identity cue manipulation, a chi-square test was used to determine whether

Table 15
Pairwise Comparisons on perceived homophily.

	Comparison	Mean	Mean difference (I-J)	Standard Error	Sig.
Low anthropomorphic visual cues	Chatbot Identity (I)	4.85	1.02	0.32	0.00
	Human identity (J)	3.83			
High anthropomorphic visual cues	Chatbot Identity (I)	3.86	−1.05	0.33	0.00
	Human identity (J)	4.91			

Table 16
Pairwise Comparisons on perceived contingency.

	Comparison	Mean	Mean difference (I-J)	Standard Error	Sig.
Low anthropomorphic visual cues	Low message (I)	4.15	−1.50	0.32	0.00
	High message (J)	5.65			
High anthropomorphic visual cues	Low message (I)	5.08	−0.26	0.33	0.43
	High message (J)	5.34			

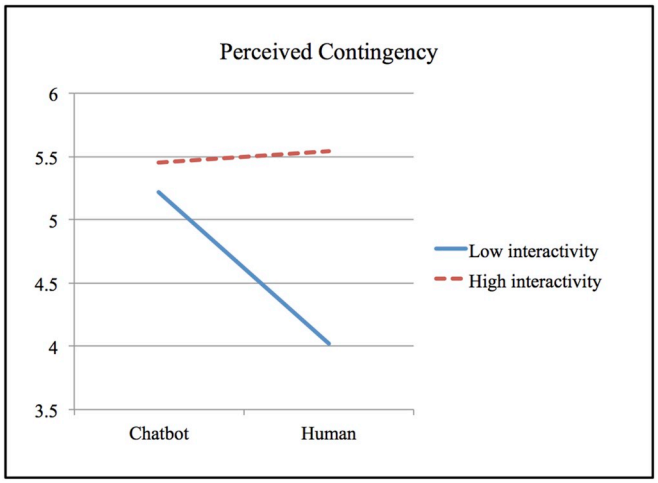

Fig. 5. Interaction effect between identity cue and message interactivity on perceived contingency.

participants perceived the agent as a human or a chat-bot. The analysis revealed that most of the participants in the human identity condition believed they had chatted with a human agent (88.7%), while most of the participants in the chat-bot condition thought they had chatted with a chat-bot (92.9%), χ^2 (1, N = 141) = 93.98, $p < .001$.

Interactivity manipulation check. Manipulation for message interactivity was also checked by performing an independent sample *t*-test, which showed that participants in the high message-interactivity condition perceived higher levels of threadedness ($M = 6.23$, $SD = 0.82$) compared to those in the low message-interactivity condition ($M = 4.82$, $SD = 1.65$), $t(139) = -6.38$, $p < .001$, Cohen's $d = 1.08$).

5.2. Hypotheses testing

The first three hypotheses predicted (1) the positive effects of the online chat agent's anthropomorphic visual cues on the degree of social presence and perceived homophily (H1), as well as (2) the mediating effects of social presence (H2) and perceived homophily (H3) on attitudes and behavioral intentions. In order to test H1, a series of three-way ANOVA tests was conducted. The tests indicated no significant main effect of the anthropomorphic visual cues on social presence, and perceived homophily (Tables 1 and 2). Therefore, H1 was not supported. To test the mediating effects of social presence (H2) and perceived homophily (H3), a bootstrapping procedure using 5000 bootstrap samples (Hayes, 2013) was performed. The results of the test found no significant mediating effect of social presence and perceived homophily on the proposed dependent variables (Tables 3 and 4). It's not surprising that any mediating effects on dependent variables were not found given that this study did not find main effects on mediators. Thus, H2 and H3 were not supported.

H4, 5, 6 predicted (1) the positive effects of the online chat agent's identity cue on the degree of social presence and perceived homophily (H4) and (2) the mediating effects of social presence (H5) and perceived homophily (H6) on attitudes and behavioral intentions. Results indicated no significant main effect of the identity cues on social presence and perceived homophily (Tables 5 and 6). Therefore, H4 was not supported. As expected, there was no significant mediating effect of social presence (Table 7) and perceived homophily (Table 8) on attitudinal and behavioral outcomes. Thus, H5 and H6 were not supported.

H7 through H11 predicted (1) the positive effects of message interactivity on social presence, perceived homophily, perceived contingency, and perceived dialogue (H7), and (2) the mediating effects of social presence (H8), perceived homophily (H9), perceived contingency (H10) and perceived dialogue (H11) on attitudinal and behavioral outcomes. As hypothesized, participants in the high message interactivity condition reported greater feelings of social presence, $F(1, 133) = 4.24$, $p < .05$, partial $\eta^2 = 0.03$, greater perception of homophily, $F(1, 133) = 4.18$, $p < .05$, partial $\eta^2 = 0.03$, greater perceived contingency, $F(1, 133) = 14.29$, $p < .001$, partial $\eta^2 = 0.10$, and greater perceived dialogue, $F(1, 133) = 19.37$, $p < .001$, partial $\eta^2 = 0.13$, compared to those in the low message interactivity condition (Tables 9 and 10). Thus, H7 was supported.

The analysis also revealed that social presence indeed mediated the effect of the agent's conversational cue on two aspects of the agent: 1) perceived expertise ($b = 0.04$, SE = 0.03, 95% C.I. from 0.00 to 0.13), 2) and perceived friendliness ($b = 0.27$, SE = 0.13, 95% C.I. from 0.03 to 0.56), as well as attitudes toward the website ($b = 0.10$, SE = 0.05, 95% C.I. from 0.02 to 0.25) and behavioral intention ($b = 0.18$, SE = 0.09, 95% C.I. from 0.03 to 0.42) (Table 11). Thus, H8 was supported.

Perceived homophily also mediated the effect of message interactivity on two aspects of the agent: 1) perceived expertise ($b = 0.06$, SE = 0.03, 95% C.I. from 0.00 to 0.17), 2) and perceived friendliness ($b = 0.10$, SE = 0.06, 95% C.I. from 0.00 to 0.28), attitudes toward the website ($b = 0.08$, SE = 0.05, 95% C.I. from 0.00 to 0.24) and behavioral intention ($b = 0.15$, SE = 0.09, 95% C.I. from 0.01 to 0.39) (Table 12). Thus, H9 was supported.

Perceived contingency was not shown as a mediator of the relationship between message interactivity and perceived expertise ($b = -0.09$, SE = 0.07, 95% C.I. from −0.27 to 0.02), perceived friendliness ($b = -0.03$, SE = 0.08, 95% C.I. from −0.19 to 0.13), attitudes toward the website ($b = -0.08$, SE = 0.08, 95% C.I. from

Table 17
Pairwise Comparisons on perceived contingency.

	Comparison	Mean	Mean difference (I-J)	Standard Error	Sig.
Low message interactivity	Chatbot Identity (I)	5.21	1.20	0.33	0.00
	Human identity (J)	4.01			
High message interactivity	Chatbot Identity (I)	5.45	−0.09	0.32	0.78
	Human identity (J)	5.54			

Table 18
Pairwise Comparisons on perceived dialogue.

	Comparison	Mean	Mean difference (I-J)	Standard Error	Sig.
Low anthropomorphic visual cues	Low message (I)	4.79	−1.22	0.26	0.00
	High message (J)	5.98			
High anthropomorphic visual cues	Low message (I)	5.48	−0.44	0.27	0.11
	High message (J)	5.92			

Table 19
Pairwise Comparisons on perceived dialogue.

	Comparison	Mean	Mean difference (I-J)	Standard Error	Sig.
Low message interactivity	Chatbot Identity (I)	5.69	1.20	0.33	0.00
	Human identity (J)	4.55			
High message interactivity	Chatbot Identity (I)	5.83	−0.09	0.33	0.78
	Human identity (J)	6.07			

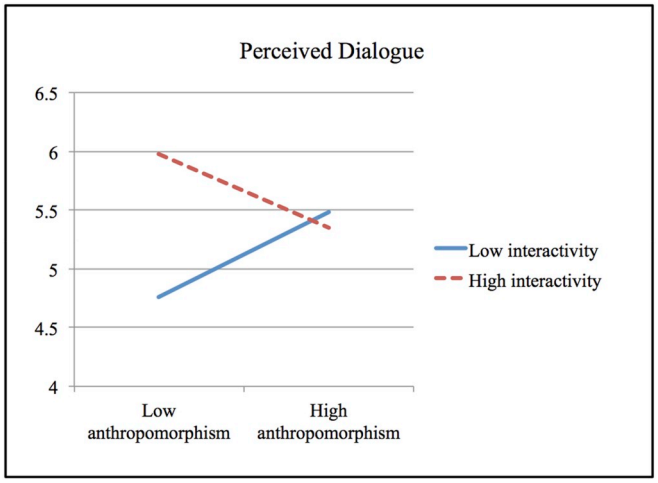

Fig. 6. Interaction effect of visual cue and message interactivity on perceived dialogue.

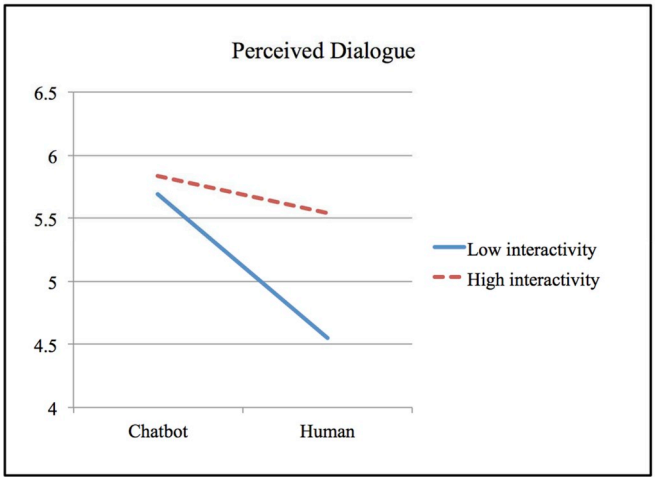

Fig. 7. Interaction effect of identity cue and message interactivity on perceived dialogue.

−0.31 to 0.06), and behavioral intention (b = −0.08, SE = 0.12, 95% C.I. from −0.37 to 0.13) (Table 13). Thus, H10 was not supported.

However, perceived dialogue was a significant mediator in the relationship between message interactivity and perceived expertise (b = 0.46, SE = 0.14, 95% C.I. from 0.23 to 0.78), perceived friendliness (b = 0.30, SE = 0.12, 95% C.I. from 0.12 to 0.61), attitudes

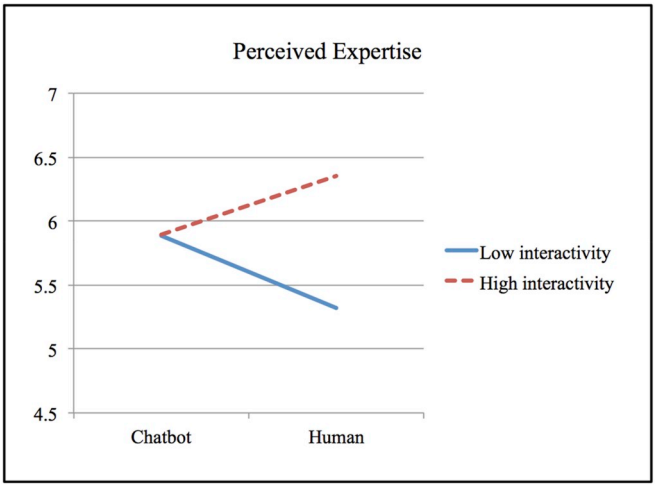

Fig. 8. Interaction effect of identity cue and message interactivity on perceived expertise of the agent.

toward the website (b = 0.59, SE = 0.18, 95% C.I. from 0.30 to 0.79), and behavioral intention (b = 0.44, SE = 0.17, 95% C.I. from 0.18 to 0.89) (Table 14). Thus, H11 was supported.

In terms of interaction effects among visual cues, identity cues, and message interactivity, first, on social presence, there was no significant interaction effect. However, on perceived homophily, there was a significant interaction effect between visual cue and identity cue, $F(1, 133) = 20.32$, $p < .001$, partial $\eta^2 = 0.13$ (Fig. 3). Follow-up tests of simple effects were used to probe the nature of the significant interaction. Results indicated when the chat agent demonstrated the low anthropmorphic visual cue, participants in the chat-bot identity condition reported greater perceptions of homophily than those in the human identity condition. On the other hand, when the chat agent had anthropomorphic visual cue, participants in the human identity condition showed greater perceived homophily than those in the chat-bot condition (Table 15). On perceived contingency, there were significant interactions between visual cue and message interactivity, $F(1, 133) = 7.08$, $p < .01$, partial $\eta^2 = 0.05$ (Fig. 4). Results of tests of simple effects (Table 16) showed that when the chat agent's visual cues were low in anthropomorphism, participants in the high message-interactivity condition perceived greater contingency than those in the low message-interactivity condition. However, when the agent's visual cues were highly anthropomorphic, there was no difference between the high message-interactivity condition and the low message-interactivity condition. There was a significant interaction between identity cue and message interactivity on perceived contingency, $F(1,$

Table 20
Pairwise Comparisons on perceived expertise of the agent.

	Comparison	Mean	Mean difference (I-J)	Standard Error	Sig.
Low message interactivity	Chatbot Identity (I)	5.88	0.55	0.21	0.01
	Human identity (J)	5.33			
High message interactivity	Chatbot Identity (I)	5.90	−0.46	0.21	0.03
	Human identity (J)	6.36			

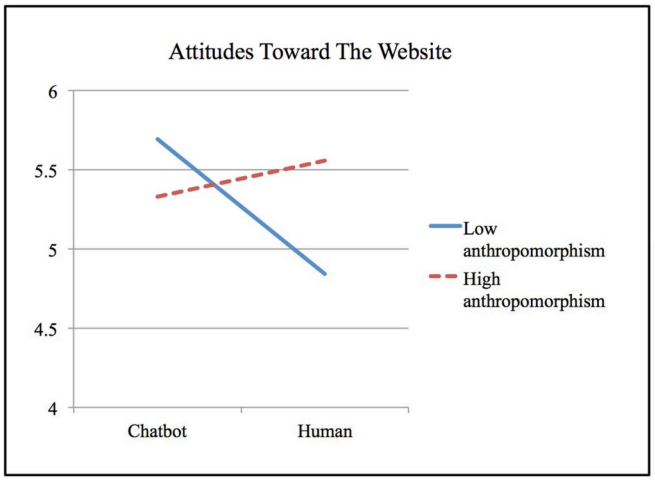

Fig. 9. Interaction effect of visual cue and identity cue on attitudes toward the website.

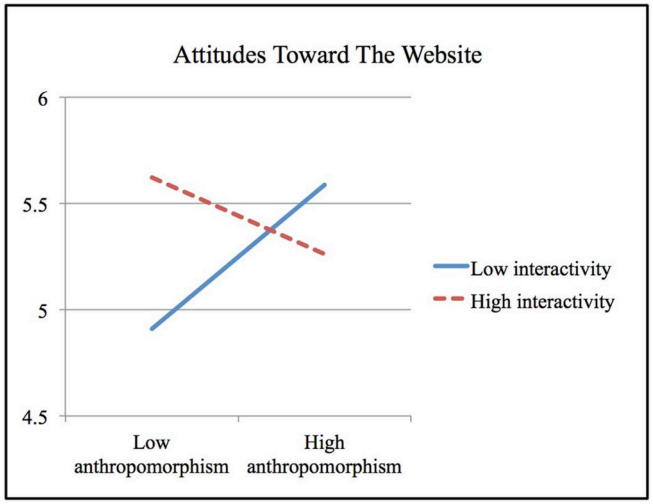

Fig. 10. Interaction effect between visual cue and message interactivity on attitudes toward the website.

133) = 7.66, $p < .01$, partial $\eta^2 = 0.05$ (Fig. 5), showing that when message interactivity was low, participants in a chat-bot condition reported greater perceived contingency than those in a human condition. However, when message interactivity was high, the identity of the chat agent did not influence the participants' perceptions of contingency (Table 17). The same patterns were shown on perceived dialogue (Tables 18 and 19) in that there was a significant interaction effect between visual cue and message interactivity, $F(1, 133) = 4.31$, $p < .05$, partial $\eta^2 = 0.03$ (Fig. 6), and between identity cue and message interactivity, $F(1, 133) = 13.47$, $p < .001$, partial $\eta^2 = 0.09$ (Fig. 7). The study also revealed an interaction effect between identity cue and message interactivity on perceived expertise of the agent, $F(1, 133) = 11.72$, $p < .001$, partial $\eta^2 = 0.08$ (Fig. 8). Tests of simple effects (Table 20) suggested that in the low message interactivity condition, the chat-bot identity led participants to perceive greater expertise of the agent than did the human identity; however, in the high message interactivity condition, the opposite pattern was found. There were also significant interactions between visual and identity cues on attitudes toward the website, $F(1, 133) = 5.22$, $p < .05$, partial $\eta^2 = 0.04$ (Fig. 9), showing that in the low anthropomorphic visual cue condition, participants perceived the identity of the chat agent to be that of a chat-bot reported more favorable attitudes toward the website than those who identified the chat agent as a human, whereas the opposite pattern was shown in the high anthropomorphic visual cue condition (Table 21). There was another significant interaction between visual cue and message interactivity, $F(1, 133) = 4.51$, $p < .05$, partial $\eta^2 = 0.03$ (Fig. 10), suggesting that when the chat agent's visual cues were low in anthropomorphism, participants in the high message-interactivity condition showed more favorable attitudes than those in the low message-interactivity condition. However, when the agent's visual cues were highly anthropomorphic, there was no difference between the high message-interactivity condition and the low message-interactivity condition (Table 22). In addition, a significant interaction between identity cue and message interactivity was found, $F(1, 133) = 4.44$, $p < .05$, partial $\eta^2 = 0.03$ (Fig. 11). Specifically, when the message interactivity was high, the agent's identity did not determine the participants' attitudes toward the website, whereas, when the message interactivity was low, participants who perceived the agent as a chat-bot reported more favorable attitudes toward the website ($M = 5.98$, $SE = 0.16$) than those who identified it as a human (Table 23). With regard to behavioral intentions, a significant interaction was found between visual and identity cues, $F(1, 133) = 5.93$, $p < .05$, partial $\eta^2 = 0.04$. As displayed in Fig. 12, the most notable difference was that between the chatbot and human identity condition among those in the low anthropomorphic visual cue condition. Follow-up tests of simple effects confirmed that this was the primary source of the interaction. (Table 24). However, there were no significant three-way interaction effects on any of the outcome variables (Table 25).

In sum, when the three humanness-related factors were deployed together on a chat-agent interface, only the main effects for message interactivity were significant. The hypothesized main effects of interface-level cues of humanness did not receive support from the data. This

Table 21
Pairwise Comparisons on attitudes toward the website.

	Comparison	Mean	Mean difference (I-J)	Standard Error	Sig.
Low anthropomorphic visual cues	Chatbot Identity (I)	5.69	0.85	0.30	0.00
	Human identity (J)	4.84			
High anthropomorphic visual cues	Chatbot Identity (I)	5.33	−0.13	0.31	0.67
	Human identity (J)	5.46			

Table 22
Pairwise Comparisons on attitudes toward the website.

	Comparison	Mean	Mean difference (I-J)	Standard Error	Sig.
Low anthropomorphic visual cues	Low message (I)	4.91	−0.71	0.30	0.02
	High message (J)	5.62			
High anthropomorphic visual cues	Low message (I)	5.50	0.20	0.31	0.51
	High message (J)	5.30			

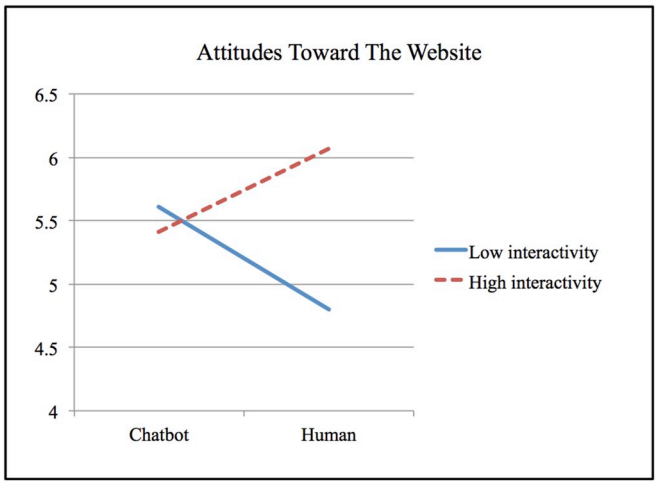

Fig. 11. Interaction effect between identity cue and message interactivity on attitudes toward the website.

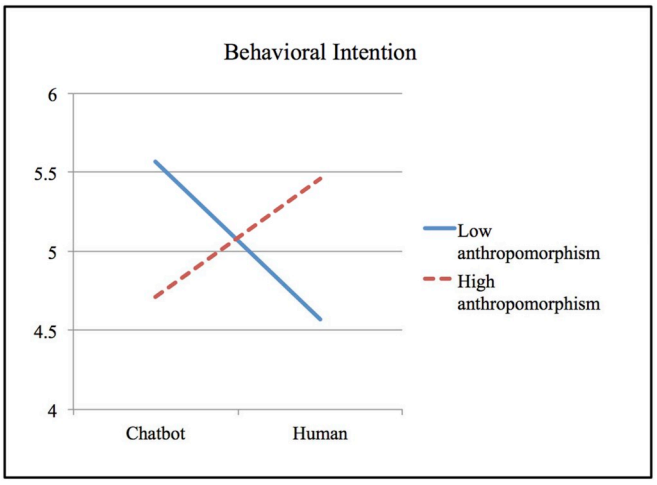

Fig. 12. Interaction effect between visual cue and identity cue on behavioral intention.

is not surprising in light of the discovery of several significant interaction effects involving these cues and message interactivity.

6. Discussion

The results of this study indicate that a chat agent's ability to deliver contingent messages (i.e., to provide message interactivity) is an important factor in determining psychological, attitudinal, and behavioral outcomes. In particular, participants in the high message interactivity condition showed greater social presence, perceived similarity, contingency, and dialogue than those in the low message interactivity condition. More importantly, a series of mediation analyses indicates that message interactivity influences participants' evaluations of the agent, attitude toward the website, and behavioral intentions by way of diverse psychological outcomes. These results particularly significant in light of the fact that the manipulation of message interactivity used in this study was quite simple—not quite as complex as real-world chatbots. That is, simple, feasible changes in their conversational styles could infuse chatbots with humanness by increasing perceptions of contingency, dialogue, social presence and homophily, which in turn contributes to positive attitudes and behavioral outcomes.

This study also discovered that message interactivity has a compensatory effect on visual interface cue and vice versa. We found that for perceived contingency, the visual anthropomorphic cues moderated the effect of message interactivity. That is, when a chat agent had a human figure icon, the effect of message interactivity was not significant, but when the agent was accompanied by a non-human icon (i.e., bubble figure), message interactivity has a positive effect on perceived contingency. Likewise, on attitudes toward the website, message interactivity moderated the effect of visual cues, such that anthropomorphic visual cues compensated for negative consequences resulting from the lack of message interactivity. From the results, we can infer that anthropomorphic visual cues help compensate for the impersonal nature of communication due to low levels of message interactivity and vice versa. In this study, users were continually exposed to the visual cue while they were chatting with agent, as the chat window showed the agent's representation (either a human figure or bubble figure) for as long as the window was open. The identity cue, on the other hand, was provided only at the beginning of the chat. This suggests that an ongoing visual cue can help compensate for low levels of humanness or impersonal communication stemming from low message interactivity while chatting.

On the other hand, different interaction patterns were found in regard to interactions between identity cues and message interactivity on perceived contingency, perceived dialogue, attitude toward the agent (expertise perception), and attitude toward the website. Specifically, when a chat agent was identified as a human, participants showed favorable evaluations if it delivered highly interactive conversation. However, when it delivered less interactive communication, participants had negative evaluations. From these findings, we can infer that the identity cue (whether a chat agent is identified at the outset as human or chatbot) seems to set expectations for the performance of the

Table 23
Pairwise Comparisons on attitudes toward the website.

	Comparison	Mean	Mean difference (I-J)	Standard Error	Sig.
Low message interactivity	Chatbot Identity (I)	5.61	0.81	0.30	0.01
	Human identity (J)	4.80			
High message interactivity	Chatbot Identity (I)	5.41	−0.09	0.30	0.76
	Human identity (J)	5.50			

Table 24
Pairwise Comparisons on behavioral intention.

	Comparison	Mean	Mean difference (I-J)	Standard Error	Sig.
Low anthropomorphic visual cues	Chatbot Identity (I)	5.57	1.00	0.36	0.01
	Human identity (J)	4.57			
High anthropomorphic visual cues	Chatbot Identity (I)	4.71	−0.25	0.37	0.48
	Human identity (J)	4.96			

Table 25
Three-way interaction effects.

Outcome variables	Identity × Visual × Message	
	F-value	Effect size
Social presence	0.01	0.00
Perceived homophily	2.81	0.02
Perceived contingency	0.00	0.00
Perceived dialogue	1.34	0.01
Attitudes toward the agent		
1) Perceived expertise	2.74	0.02
2) Perceived friendliness	1.61	0.01
Attitudes toward the website	0.79	0.01
Behavioral intention	2.09	0.01

agent. As the MAIN model suggests, a human identity cue may trigger a cognitive heuristic about the essential characteristics of humanness. This subsequently leads individuals to expect more contingent conversations with a chat agent, as contingency is a core characteristic of human-to-human communication. Such high expectations extend to human agents online: indeed, in these cases, participants' expectations cause them to evaluate human agent's low contingent conversational style much more negatively than they otherwise might. Given that the negative violation associated with the human identity cue condition was more salient in the results, it may be speculated that individuals have much clearer heuristics regarding human-associated characteristics than chat-bot- or machine-associated characteristics.

Lastly, this study also found interaction effect between identity and visual cues. Specifically, when the chat agent was represented with a less anthropomorphic visual cue, participants in the human identity condition reported lower perceptions of homophily, less favorable attitudes toward the website, and lower behavioral intentions than those in the chatbot identity condition. On the other hand, when the chat agent displayed a highly anthropomorphic visual cue, participants in the chatbot identity condition showed lower perceptions of homophily, less favorable attitudes, and lower behavioral intentions than those in the human identity condition. This finding suggests another kind of expectancy violation effect because identifying the agent as a human may have raised participants' expectations for humanness, only to be violated by the use of a low-anthropomorphic cue (like the picture of a dialog bubble) and resulting in lower ratings of homophily, attitudes, and behavioral intention. And, identifying the agent as a chatbot may have made them expect machine-like cues, but the use of a highly anthropomorphic visual cue may have been perceived as disingenuous and thereby reduced feelings of homophily, favorable attitudes, and behavioral intentions. This further supports the notion that identity cues, in the form of labeling a machine agent, can be quite influential in framing the interaction and thereby shaping user perceptions, including perceived similarity and affiliative tendency.

The expectancy violation effect has been explored in many studies across several domains (Burgoon & Walther, 1990; Cappella & Greene, 1982; Kahneman & Miller, 1986). The one discovered in this study can have important consequences for the nature and success of human interactions with chatbots. Combined with the earlier interaction finding, it appears that triggering the “machine heuristic” (Sundar, 2008) with the help of identity cues may lower user expectations for humanlike conversations and therefore be perceived more positively, as long those visual cues on the interface are low on anthropomorphism. While highly anthropomorphic visual cues can compensate for low message interactivity, the identity cue cannot do the same. This cue has primacy over all other cues because it signals the identity of the source. As scholars of human-computer interaction have long demonstrated, “source orientation” is a critical determinant of user experience. Competitions like the Turing Test, which inspire designs of systems that can fool human users into believing they are humans, are the holy grail precisely because they capitalize on differential psychological predispositions toward human vs. machine sources. While humanlikeness is the ultimate goal, most users are impressed when a machine approximates humanlike appearance or behavior to a certain degree and generally have lower expectations for message interactivity with machines. Our study shows that revealing the identity of the machine can capitalize on these expectations and lead to more satisfactory interactions with chatbots.

This study has valuable practical implications for interface and interaction designers seeking to incorporate or improve the live-chatting function within their online media platforms. First, the effectiveness of our manipulations provide simple operationalizations of an online agent's visual cues, identity cue, and message interactivity, paving the way for designers and researchers to imbue humanness in chat-bots. The results also point to the need for organizational caution when employing chat-bots. Given the series of interaction effects documented here, especially the expectancy violation effects, it is clear that users tend to have different expectations of agents' performances depending on whether the agents are identified as humans or chat-bots. Therefore, agents' actual communication styles should align with users' expectations. The results also suggest that the best way to compensate for the lack of face-to-face communication might be to make the agent's communication as contingent as possible.

6.1. Limitation and future study

This study also has several limitations. The sample from Amazon's Mechanical Turk, while yielding more generalizable and reliable results than samples comprised of college students (Buhrmester et al., 2011; Goodman et al., 2013), requires the study to be conducted online, with each participant completing the study in his or her own preferred environment, thus introducing noise into the study's procedures. To simulate the experience of chatting and to control for content effects, thereby identifying the effects of different levels of message contingency, this study used a scenario that asked participants to follow a certain chatting procedure. While the use of such a scenario allowed the study to control unwanted effects that may emerge in an online experiment, it also introduced low ecological validity into the study. Future research would do well to address these limitations.

6.2. Conclusion

In conclusion, the current study contributes to researchers' knowledge about three humanness-associated factors of online chat agents and how the interactions among these three factors influence psychological, attitudinal, and behavioral outcomes. This study showed that message interactivity works to compensate for the impersonal nature associated with low anthropomorphic visual cue and chat-bot identity

cue. Moreover, the identity cue turned out to be a key factor in eliciting certain expectations regarding the agent's performance in conversation. Therefore, depending on the identity cue, the effects of message interactivity, or the agent's performance, may vary. Given the theoretical complexity of individuals' psychological, attitudinal, and behavioral responses to the agent as discovered in this study, this study suggests that website designers and public relations executives carefully implement the live-chatting function using chat agents in order to help mitigate users' negative evaluations or experiences due to high expectations.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.chb.2019.01.020.

References

Ashforth, B. E., & Humphrey, R. H. (1997). The ubiquity and potency of labeling in organizations. *Organization Science, 8*(1), 43–58.

Baylor, A., & Kim, Y. (2003). The role of gender and ethnicity in pedagogical agent perception. *World conference on E-learning in corporate, government, healthcare, and higher education: Vol. 2003*, (pp. 1503–1506). No. 1.

Bellur, S., & Sundar, S. S. (2017). Talking health with a machine: How does message interactivity affect attitudes and cognitions? *Human Communication Research, 43*(1), 25–43. https://doi.org/10.1111/hcre.12094.

Bente, G., Rüggenberg, S., Krämer, N. C., & Eschenburg, F. (2008). Avatar-mediated networking: Increasing social presence and interpersonal trust in net-based collaborations. *Human Communication Research, 34*(2), 287–318. https://doi.org/10.1111/j.1468-2958.2008.00322.x.

Biocca, F., Harms, C., & Burgoon, J. (2003). Toward a more robust theory and measure of social presence: Review and suggested criteria. *Presence, 12*(5), 456–480. https://doi.org/10.1162/105474603322761270.

Buhrmester, M., Kwang, T., & Gosling, S. D. (2011). Amazon's Mechanical Turk a new source of inexpensive, yet high-quality, data? *Perspectives on Psychological Science, 6*(1), 3–5. https://doi.org/10.1177/1745691610393980.

Burgoon, J. K., & Walther, J. B. (1990). Nonverbal expectancies and the evaluative consequences of violations. *Human Communication Research, 17*(2), 232–265. https://doi.org/10.1111/j.1468-2958.1990.tb00232.x.

Cappella, J. N., & Greene, J. O. (1982). A discrepancy-arousal explanation of mutual influence in expressive behavior for adult and infant-adult interaction. *Communication Monographs, 49*(2), 89–114. https://doi.org/10.1080/03637758209376074.

Cyr, D., Head, M., Larios, H., & Pan, B. (2009). Exploring human images in website design: A multi-method approach. *MIS Quarterly, 33*(3), 539–566. https://doi.org/10.2307/20650308.

Gefen, D., & Straub, D. W. (2003). Managing user trust in B2C e-services. *e-Service Journal, 2*(2), 7–24. https://doi.org/10.2979/esj.2003.2.2.7.

Gelman, S. A., & Heyman, G. D. (1999). Carrot-eaters and creature-believers: The effects of lexicalization on children's inferences about social categories. *Psychological Science, 10*(6), 489–493. https://doi.org/10.1111/1467-9280.00194.

Ghose, S., & Barua, J. J. (2013, May). Toward the implementation of a topic specific dialogue based natural language chatbot as an undergraduate advisor. *Informatics, electronics & vision (ICIEV), 2013 international conference on* (pp. 1–5). IEEE.

Goethals, G. R., & Nelson, R. E. (1973). Similarity in the influence process: The belief-value distinction. *Journal of Personality and Social Psychology, 25*(1), 117. https://doi.org/10.1037/h0034266.

Gong, L., & Nass, C. (2007). When a talking-face computer agent is half-human and half-humanoid: Human identity and consistency preference. *Human Communication Research, 33*(2), 163–193. https://doi.org/10.1111/j.1468-2958.2007.00295.x.

Goodman, J. K., Cryder, C. E., & Cheema, A. (2013). Data collection in a flat world: The strengths and weaknesses of Mechanical Turk samples. *Journal of Behavioral Decision Making, 26*(3), 213–224. https://doi.org/10.1002/bdm.1753.

Gosling, S. D., Vazire, S., Srivastava, S., & John, O. P. (2004). Should we trust web-based studies? A comparative analysis of six preconceptions about internet questionnaires. *American Psychologist, 59*(2), 93. https://doi.org/10.1037/0003-066X.59.2.93.

Hayes, A. F. (2013). *An introduction to mediation, moderation, and conditional process analysis.* New York: The Guilford Press.

Hrastinski, S. (2008). Asynchronous and synchronous e-learning. *Educause Quarterly, 31*(4), 51–55.

Hu, Y., & Sundar, S. S. (2010). Effects of online health sources on credibility and behavioral intentions. *Communication Research, 37*(1), 105–132. https://doi.org/10.1177/0093650209351512.

Kahneman, D., & Miller, D. T. (1986). Norm theory: Comparing reality to its alternatives. *Psychological Review, 93*(2), 136–153. https://doi.org/10.1037/0033-295X.93.2.136.

Kent, M. L., & Taylor, M. (2002). Toward a dialogic theory of public relations. *Public Relations Review, 28*(1), 21–37. https://doi.org/10.1016/S0363-8111(02)00108-X.

Kim, J. (2010). Balancing uniqueness and assimilation in computer-mediated groups. *Computers in Human Behavior, 26*(4), 778–784. https://doi.org/10.1016/j.chb.2010.02.001.

Kim, Y., & Sundar, S. S. (2012). Anthropomorphism of computers: Is it mindful or mindless? *Computers in Human Behavior, 28*(1), 241–250. https://doi.org/10.1016/j.chb.2011.09.006.

Koda, T. (1996). *Agents with faces: A study on the effects of personification of software agents* (Unpublished master thesis). Boston, MA, USA: Massachusetts Institute of Technology.

Koh, Y. J., & Sundar, S. S. (2010). Heuristic versus systematic processing of specialist versus generalist sources in online media. *Human Communication Research, 36*(2), 103–124. https://doi.org/10.1111/j.1468-2958.2010.01370.x.

Lee, E.-J., & Oh, S. Y. (2015). Effects of visual cues on social perceptions and self-categorization in computer-mediated communication. In S. S. Sundar (Ed.). *The handbook of the psychology of communication technology* (pp. 115–136). Malden, MA: Wiley Blackwell.

Mason, W., & Suri, S. (2012). Conducting behavioral research on Amazon's Mechanical Turk. *Behavior Research Methods, 44*(1), 1–23.

Mason, W. A., & Watts, D. J. (2009). Financial incentives and the performance of crowds. *Proceedings of the ACM SIGKDD workshop on human computation* (pp. 77–85). New York: ACM.

Nowak, K. L. (2004). The influence of anthropomorphism and agency on social judgment in online chat environments. *Journal of Computer-Mediated Communication, 9*(2), https://doi.org/10.1111/j.1083-6101.2004.tb00284.x.

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), 153–178. https://doi.org/10.1111/j.1083-6101.2006.tb00308.x.

Palmer, M. (1995). Interpersonal communication and virtual reality: Mediating interpersonal relationships. In F. Biocca, & M. Levy (Eds.). *Communication in the age of virtual reality* (pp. 277–299). Mahwah, NJ: Lawrence Erlbaum Associates.

Rafaeli, S. (1988). Interactivity: From new media to communication. In R. P. Hawkins, J. M. Wiemann, & S. Pingree (Eds.). *Advancing communication science: Merging mass and interpersonal processes* (pp. 110–134). Newbury Park, CA: Sage.

Rice, R. E. (1984). Mediated group communication. In R. E. Rice (Ed.). *The new media: Communication, research, and technology* (pp. 129–154). Beverly Hills, CA: Sage.

Rocca, K. A., & McCroskey, J. C. (1999). The interrelationship of student ratings of instructors' immediacy, verbal aggressiveness, homophily, and interpersonal attraction. *Communication Education, 48*(4), 308–316. https://doi.org/10.1080/03634529909379181.

Rogers, E. M., & Bhowmik, D. K. (1970). Homophily-heterophily: Relational concepts for communication research. *Public Opinion Quarterly, 34*(4), 523–538. https://doi.org/10.1086/267838.

Seltzer, T., & Mitrook, M. A. (2007). The dialogic potential of weblogs in relationship building. *Public Relations Review, 33*(2), 227–229. https://doi.org/10.1016/j.pubrev.2007.02.011.

Short, J., Williams, E., & Christie, B. (1976). *The social psychology of telecommunications.* New York, NY: John Wiley & Sons.

Sundar, S. S. (2008). The MAIN model: A heuristic approach to understanding technology effects on credibility. In M. J. Metzger, & A. J. Flanagin (Eds.). *Digital, media, youth, and credibility* (pp. 72–100). Cambridge, MA: The MIT Press. https://doi.org/10.1162/dmal.9780262562324.073.

Sundar, S. S. (2009). Media effects 2.0: Social and psychological effects of communication technologies. In R. L. Nabi, & M. B. Oliver (Eds.). *The SAGE handbook of media processes and effects* (pp. 545–560). Thousand Oaks, CA: Sage Publications.

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

Sundar, S. S., Go, E., Kim, H. S., & Zhang, B. (2015). Communicating art, virtually! Psychological effects of technological affordances in a virtual museum. *International Journal of Human-Computer Interaction, 31*, 385–401. https://doi.org/10.1080/10447318.2015.1033912.

Sundar, S. S., Oeldorf-Hirsch, A., Nussbaum, J. F., & Behr, R. A. (2011). Retirees on Facebook: Can online social networking enhance their health and wellness? *Proceedings of the 2011 annual conference extended abstracts on human factors in computing systems (CHI EA'11)* (pp. 2287–2292). .

Sundar, S. S., Xu, Q., & Bellur, S. (2010, April). Designing interactivity in media interfaces: A communications perspective. *Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 2247–2256). ACM.

Suri, S., & Watts, D. J. (2011). Cooperation and contagion in Web-based, networked public goods experiments. *PLoS One, 6*(3) e16836.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185*(4157), 1124–1131.

Wexelblat, A. (1998). Don't make that face: A report on anthropomorphizing an interface. *Intelligent Environments, 173*(179) 98-02.