Evaluating Anthropomorphic Product Recommendation Agents: A Social Relationship Perspective to Designing Information Systems

LINGYUN QIU AND IZAK BENBASAT

LINGYUN QIU is an Assistant Professor in the Department of Management Science and Information Systems at the Guanghua School of Management, Peking University, China. He received his Ph.D. in Management Information Systems from the University of British Columbia, Canada. His research interests include human–computer interactions, social interactions in electronic commerce, and decision support systems.

IZAK BENBASAT is a Fellow of the Royal Society of Canada and a CANADA Research Chair in Information Technology Management at the Sauder School of Business, University of British Columbia, Canada. He received his Ph.D. in Management Information Systems from the University of Minnesota. The general theme of his research is improving the interactions among IT, management, and IT users. He currently serves on the board of the *Journal of Management Information Systems.* He was editor-in-chief of *Information Systems Research,* editor of the Information Systems and Decision Support Systems Department of *Management Science,* and a senior editor of *MIS Quarterly.*

ABSTRACT: In online shopping environments, the product-advising function originally performed by salespeople is being increasingly taken over by software-based product recommendation agents (PRAs). However, the literature has mostly focused on the functionality design and utilitarian value of such decision support systems, mostly ignoring the potential social influence they could exert on their users.

The objective of this study is to apply a *social relationship* perspective to the design of interfaces for PRAs. We investigate the effects of applying anthropomorphic interfaces—namely, humanoid embodiment and voice output—on users' perceived social relationship with a technological and software-based artifact designed for electronic commerce contexts. The findings from a laboratory experiment indicate that using humanoid embodiment and human voice–based communication significantly influences users' perceptions of social presence, which in turn enhances users' trusting beliefs, perceptions of enjoyment, and ultimately, their intentions to use the agent as a decision aid. These results extend the applicability of theories concerning traditional shopper–salesperson relationships to customers' interactions with technological artifacts residing on Web sites—that is, the recommendation agent software—and provide practitioners with guidelines on how to design Internet stores with the goal of building social relationships with online shoppers and enhancing their overall shopping experiences.

Journal of Management Information Systems / Spring 2009, Vol. 25, No. 4, pp. 145–181.

0742–1222 / 2009 $9.50 + 0.00.
DOI 10.2753/MIS0742-1222250405

KEY WORDS AND PHRASES: avatar, decision support systems, electronic commerce, human voice, laboratory experiments, recommendation agents, social presence, text-to-speech (TTS), user acceptance of IT.

IN ONLINE SHOPPING ENVIRONMENTS, the product-advising function originally performed by salespeople is being increasingly taken over by software-based product recommendation agents (PRAs). PRAs help to alleviate consumers' cognitive load of gathering, screening, and evaluating vast amounts of product information available on the Web by eliciting customers' needs and then identifying the products that best fit them [127]. By implementing these types of software agents, comparison shopping Web sites, such as mySimon.com, Shopping.com, and MyProductAdvisor.com, are getting increasingly popular among online shoppers who rely on the Internet to search for product information and evaluate a large number of alternatives [19]. eBay's acquisition in 2005 of Shopping.com for $620 million further attests to the importance of recommendation technologies to e-commerce leaders. Empirical evaluations of PRAs have so far focused on their *utilitarian* values (see [127] for a review of the PRA literature). Using a transactional perspective, these studies have investigated issues such as the extent to which a PRA saves consumers cognitive effort and improves the quality of their purchase decisions (e.g., [45, 99]).

In an attempt to adapt the traditional theories of information technology (IT) use and adoption to better explain users' behaviors in the e-commerce context, social, relational, and affective constructs have been included in adoption models, complementing cognitive beliefs, to explain users' adoption of online shopping [38, 59, 60, 120]. However, the social and relationship-building aspects of *consumer–PRA interactions* have gained attention only recently [2, 58], in which users' perceptions of their relationship with a PRA are assessed by the construct *social presence.*

The term *social presence* was originally developed to measure people's subjective perceptions of other people when the communication is mediated by communication medium [112]. It describes the extent to which a medium is perceived as sociable, warm, sensitive, personal, or intimate when it is used to interact with others [74]. Although social presence research has focused primarily on people's connections to other *human beings* who are mediated (linked) by technology, recent studies have argued that the concept of social presence is equally applicable to people's interactions with technological artifacts as well, such as computers, machines, or software agents [17, 18, 88, 94]. Particularly, Nass and Moon, through a series of experiments, have demonstrated that individuals would mindlessly apply social rules and expectations to computers [88]. In fact, people's social responses could be automatically elicited by any cues that are related to human characteristics [104]. The term *parasocial interaction* is also used in the media and communication literature to describe the psychology phenomenon that media users tend to develop mental connections with the characters presented on the media [40, 43].

In the context of online shopping, social presence has also been employed to evaluate the social aspects of the relationship between users and *Web sites* [38, 44, 60]. Researchers have investigated various Web site characteristics that could enhance users' perceptions of social presence, such as socially rich text contents and personalized greetings [38], emotive text and pictures of humans [44], and functionalities such as live chat and online reviews [27]. However, few studies have explored particular IT artifacts design that could be applied to strengthen the perceived social presence of a PRA.

The importance of design is well recognized in information systems (IS) research [48]. As Benbasat and Zmud [13, 14] have argued, the relevance of IS research is directly related to its applicability to design and implementation for IS practice. Therefore, not only should researchers underline and explicate how a PRA's perceived social presence could benefit the online retailer's bottom line, it is also imperative for them to propose and examine various alternatives that could be used by IT professionals for the *design* of social connections between users and PRAs. Although designing useful artifacts is a complex endeavor due to the fast advances in technical knowledge [48], it is important to look into the potentials of emerging IT artifacts in new application areas. As one such investigation, this study proposes and examines two anthropomorphic[1] features that could be applied to PRA interfaces—namely, humanoid embodiment and voice output.

The word *agent* conventionally has referred to a human being (e.g., a real estate broker) who is empowered to provide guidance. Advice-giving systems designed for online shopping environments are also likened to people, such as a *super salesperson* [1] or a *wise uncle* [108]. For this reason, a seemingly natural approach of fulfilling the metaphor of a sociable and interpersonal consultation experience is to endow the PRA with an anthropomorphic interface [7]. Nevertheless, as the historic debates in the area of human–computer interaction on the appropriateness of embodying computers with human-like interfaces (see [30] for a review) indicate, the efficacy of such a design approach is by no means universally accepted. Consequently, there is a need to explore such issues further.

This study has two major goals: (1) to investigate how IT design can influence the perceptions of social presence between users and PRA by embodying PRAs with humanoid embodiment and speech output, and (2) to further examine how social presence influences the antecedents to consumers' intentions to use the PRA in the future as a decision aid.

Literature Review

Anthropomorphic PRA Interfaces

ACADEMIC RESEARCHERS HAVE BEEN ATTEMPTING to replicate the influence of nonverbal cues in interpersonal communications to human–computer interactions by deploying anthropomorphic interfaces to software agents. Examples include the real estate agent "REA" [16, 21], the 3D realistic character "Greta" that serves as an advice-giving

agent [15, 32], and agents for pedagogical purposes [69, 70]. There are also a growing number of commercial products and Web-based services that enable designers to create a wide array of animated speaking characters that can be added to Web sites for a variety of professional and personal applications, such as those developed by Oddcast (www.oddcast.com) and Haptek (www.haptek.com). These systems vary greatly in their linguistic capabilities, output modalities, and task domains, but all share common features in that they attempt to engage users in a natural conversation with an on-screen face or body.

The majority of previous anthropomorphic interface research has focused on developing and examining various designs of an agent's appearance and behavior, such as avatar dimensionality (i.e., 2D or 3D) [77], facial expressions [9, 15, 73], communication modalities [82], and nonverbal behavior [26]. These designs have also been evaluated in various application contexts, such as computer-aided learning [31] and electronic commerce [49, 55, 77]. A summary of empirical evaluations of anthropomorphic interfaces is provided in Appendix A.

Recent research on anthropomorphic interface agents has shown that multimodal information channels of an embodied agent, e.g., the capabilities of exchanging information via voice, gestures, gaze, and facial expressions in addition to language, have the potential of providing users with natural means to foster affinity, trust, and pleasure as well as to exchange information with computers [21, 22]. Nevertheless, as illustrated in Appendix A, there is a paucity of studies of anthropomorphic interfaces in online shopping contexts; hence, the effectiveness of such interfaces for product recommendation systems has yet to be empirically investigated. This paucity coincides with the fact that such designs have not yet been widely utilized by online vendors or comparison shopping Web sites, which could be partially explained by the technological constraints faced by most shopping Web sites. For example, security concerns might discourage users from adopting such designs: in order to have a fully embodied and animated character visualized by Internet browsers, users may have to download and install additional software (also known as "plug-in" programs), which might induce concerns for virus or malicious software, particularly for unfamiliar e-commerce Web sites. More importantly, most managers may still doubt that the extra cost for designing and maintaining highly engaging anthropomorphic interfaces is financially justifiable, especially in the absence of evidence to confirm that such designs do in fact contribute to favorable customer experiences and, consequently, the Web site's conversion rate.

Among the few studies that evaluated the application of anthropomorphism in online shopping, McBreen and Jack [77] examined various implementations of humanoid synthetic agents, such as 3D talking heads, 2D images, or disembodied voice. However, participants of this study did *not personally interact* with such agents; instead, they were asked to view two-minute videos showing the dialogue between another "customer" and a synthetic agent. Due to the lack of intrinsic involvement and the absence of actual interactions, the subjects are unlikely to experience the subtle social relationship–building process with a PRA, thus perceptions based on such a "third party" perspective may be inadequate. Keeling et al. [55] interviewed 30 Internet shop-

pers about the acceptability and viability of implementing embodied conversational agents on retail Web sites. However, these shoppers were only asked to observe some real-world embodied agents designed for *nonretailing Web sites* and then comment on "what if these agents were used in online shopping Web sites." This design may have also undermined the validity of the results because people may have very different experiences when actually interacting with an agent to perform a real shopping task. In a recent study by Holzwarth et al. [49], participants were asked to interact with and evaluate humanoid agents designed to deliver product information in an online shopping Web site. However, these agents only helped shoppers *customize* a product rather than recommended what to buy. As such, it is significantly different from PRAs on most shopping Web sites that help shoppers filter a myriad of product attributes and provide a shortlist of those conforming to shoppers' expressed preferences.

Another major limitation of the aforementioned studies is that they either have not studied the effects of speech output or have employed designs that confounded the effects of humanoid embodiments with those of multimodal (visual and audio) presentations. For online vendors, having a PRA with a voice interface incurs extra costs compared to using on-screen text. Also, vendors have to decide on a speech interface implemented either with computer-synthesized voice or prerecorded human speech, as cost and voice quality vary greatly between these two.

Human–PRA Social Interactions and Agent Usage Intentions

According to the "relationship marketing" literature, building social relationships between salespeople and customers increases customers' satisfaction, loyalty, favorable word-of-mouth recommendations, and purchases. Consumers in turn receive *functional* and *social* benefits [105]. Functional benefits include time savings, convenience, fashion advice, and better purchase decisions, whereas social benefits include enjoying the salesperson's company and spending time with the salesperson.

In the context of e-commerce, PRAs assume the role of product advising typically associated with salespeople in physical settings. As PRAs are basically decision support systems, users' intentions to use them could be partly explained by traditional models of IS adoption (e.g., the technology acceptance model [TAM] [28]). However, the cognitive-dominated beliefs, such as *perceived usefulness* and *perceived ease of use* in TAM, fail to capture the social benefits PRAs offer. In fact, technological artifacts in e-commerce settings (e.g., PRAs) are perceived by their users as both interaction mediators and social actors [2, 58]. This is explained by the "computers are social actors" (CASA) paradigm [104], which posits that people ascribe social attributes to computer technology, especially when it is perceived to possess characteristics normally associated with human behavior. When a PRA gives highly customized recommendations in response to users' needs, people may even develop social and emotional bonds with the agent [58].

In this research, we investigate *social presence,* a construct that refers to the feeling of "being with another" [18], to describe the "quasi-social" relationships between the agent and its users and to evaluate the users' perceptual differences of the agent's social

characteristics which is likely to be influenced by the anthropomorphic interface components. Earlier studies of social presence have generally investigated technologically mediated representations of other human beings, while recent ones have extended the boundaries to artificial representations of humanoid or animal-like intelligence, such as virtual human agents, computers, and robots. As noted by Lee and Nass [64], social presence can be created through agents or through the machine itself, especially when the way people make sense of virtually presented others is undergoing significant changes with more frequent experiences in mediated virtual environments (via agents and avatars) and increasing communications among projected selves.

Online shoppers' perceptions of social presence have been found to influence online consumer trust [38, 44] and contribute to customer loyalty in e-service environments [27, 60], as well as positively correlate with TAM-related constructs, such as perceived usefulness and perceived enjoyment [27, 44]. However, the effects of social presence on *PRA usage intentions* have not been empirically investigated.

Hypothesis Development

To illustrate how anthropomorphic interfaces will ultimately impact consumers' PRA usage intentions by virtue of first influencing social presence, we utilize the research model shown in Figure 1. It should be noted that Cyr et al. [27] previously proposed a similar model and examined the direct effect of social presence on consumer's loyalty to online vendors, as well as its indirect effects mediated by trust, perceived usefulness, and perceived enjoyment.[2] However, though the same set of mediating variables is investigated in this study, there exist significant differences in the application context (PRAs versus online shopping Web sites), the dependent variable (agent usage intentions versus consumer loyalty to the online vendor), and the IT artifacts utilized to effect social presence. The constructs of the model and their relationships are discussed below.

PRA Interface and Social Presence

Although people to some degree respond socially to computers, whether they have humanoid embodiments or not [56], their social responses appear to be stronger with more "human"-looking images [92, 118]. This is the *persona effect* [70] explained by the fact that people rely on physical characteristics of their communication partners as the basis for identification [3, 50]. Investigations of implementing human faces in online settings have found that people perceive social interconnections with, and attribute personality traits to, human faces they see on a computer. These anthropomorphic representations could be either photographs of other people [44, 106] or a variety of software-generated "talking faces" [68, 113, 122]. More specifically, a PRA with humanoid embodiment can engage the user in interactions that are more similar to face-to-face conversations, as users are provided with more visual signals from the agent. In particular, the animated face may exhibit a number of nonverbal cues, such

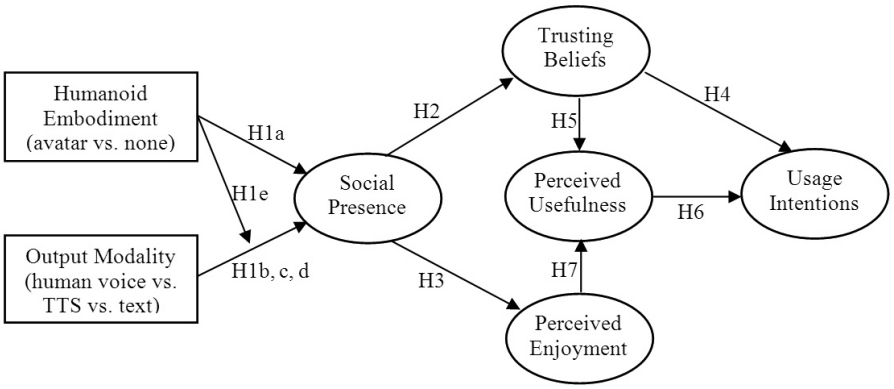

Figure 1. The Research Model

as eye contact, gazing, and expressive facial expressions such as lip movements and smiles, all of which contribute to feelings of social interactions.

Another source of social presence is reciprocity [63]. When interacting with human-like virtual characters, people reciprocate first before thinking about the ontological nature of the virtual actors. The *social agency theory* proposed by Mayer et al. [76] has argued that social cues in a multimedia message can prime the social conversation schema in receivers. Once this schema is activated, receivers are more likely to act as if they are in a conversation with another person. When a PRA is embodied with an animated face, this observable human characteristic conveys apparent cues of interpersonal communication and the user reacts by applying social norms reciprocally. The presence of anthropomorphic features impedes users from carefully elaborating the ontological nature of the virtual agent. Instead, they may judge, evaluate, and respond to the animated character by subconsciously applying the social rules practiced in everyday life. Hence, we expect that users interacting with an embodied PRA will be more likely to treat the interaction as an interpersonal conversation.

> *Hypothesis 1a: Users will perceive greater social presence when interacting with a PRA with a humanoid embodiment compared to the same PRA without a humanoid embodiment.*

Communication by voice has long been found to be able to generate physiological and affective arousal [110, 111], as voice could carry much richer information than the literal meaning of its denotative content; for example, people can detect differences between emotions based on the tone of a voice almost as well as from facial expressions [111]; people can also perceive social status, personality, and attitudes from others' speech styles [42], and use voice communication to determine group identity characteristics, such as age, gender, geographic origins [101], as well as emotions [111] and attitudes [41]. Studies on computer-mediated communications have repeatedly revealed that perceived social presence is higher in the audio-conference condition than text-chat condition [96, 109].

The use of speech on software interfaces can significantly strengthen the interpersonal nature of human–computer interactions. Voice is powerful indicator of presence of another person [85] and promotes richer processing by incorporating additional nonverbal cues through varying tones, speed, and intonation. Research has revealed that users make attributions about the interfaces with voice output using the same rules and heuristics they normally apply to humans [89]. For example, a single computer with a variety of voice identities could elicit perceptions of multiple distinct entities [90] or evoke gender stereotypes [93]. Therefore, we posit:

Hypothesis 1b: Users will perceive greater social presence when interacting with a PRA with human speech output compared to the same PRA with textual output.

Computer-synthesized speech generators, also known as text-to-speech (TTS) systems, are designed to read text aloud [33]. Compared with prerecorded human speech, TTS systems enjoy a high level of flexibility at much lower operation cost, inasmuch as they could generate speech on-the-fly and create almost infinite arrays of content. Nevertheless, currently TTS has some inherent deficiencies because it lacks the quality and prosody of natural human speech. TTS voices tend to exhibit unnatural pauses, accents, and word emphases, as well as discontinuities between phonemes and syllables [87]. Consequently, most TTS systems sound somewhat "nonhuman" [89], which restricts the social cues they can convey [76]. As a result, we expect that

Hypothesis 1c: Users will perceive greater social presence when interacting with a PRA with human speech output compared to the same PRA with TTS output.

On the other hand, a series of studies by Lee [61], Lee et al. [62], and Nass and Lee [87] have also revealed that people still attribute some social characteristics to computer-generated voices. They attempt to perceive social cues embedded in voices even when they know that the voice is generated by software [61, 65, 87]. Limited prosodic features, such as intonation and phrasing, are still discernible by people who are primed to the distinctive social cues embedded in human speech, such as gender and emotional status, to the extent that listeners unconsciously attribute some of these cues to computer-generated speech. Empirical evidence indicates that people exhibit similarity-attraction toward computer-generated speech [87]. Therefore,

Hypothesis 1d: Users will perceive greater social presence when interacting with a PRA with TTS output compared to the same PRA with textual output.

Although both humanoid embodiment and output modality are expected to have distinct influences on users' perceptions of social presence, prior research has also suggested that these two factors could be somewhat interrelated. The investigations on potential interactions between the two interface components are based on the perceived consistency theories. Two types of consistency have received increasing attention in both social psychology and, recently, human–computer interaction (HCI) literature (e.g., [52]): consistency across time (temporal consistency) and consistency across components (component consistency). While *temporal consistency* refers to an interaction partner's characteristics remaining consistent throughout the relationship,

component consistency is the one of interest in this study because it is related to the match between the presence of humanoid embodiments and output modalities.

The *consistency* across interface components could influence users' evaluation above and beyond the individual effect of each element. People prefer to interact with those who are consistent in their characteristics, because consistency in others lightens one's cognitive load and makes it easier to predict what will happen once engaged with others [34]. This happens even when they are interacting with computers. For example, Nass and Gong [84] have found that when the perceived humanness was *matched* across both the visual and voice modalities, individuals were more willing to disclose personal information and felt more comfortable using the interface. In other words, a more thoroughly "humanized" interface will be better perceived than a partial one. Study in the domain of virtual reality has also revealed that feeling of presence was low when there was a large mismatch between the appearance and behavioral realism of an embodied agent [5]. The *consistency* between a PRA's interface characteristics makes it easier for PRA users to suspend their disbelief in its "humanness," hence making them more deeply engrossed in the virtual environment and, consequently, enhancing their feelings of social presence. Therefore,

> *Hypothesis 1e: The modality effect (i.e., human speech versus textual output) concerning social presence will be more significant for the PRA with a humanoid embodiment than the PRA not so embodied.*

Social Presence and Usage Intentions

Social Presence and Trust

Trust in PRA is considered as an extension of interpersonal trust. When consumers form their initial trusting beliefs in a PRA, the perceived quality of the information provided by the agent contributes to the cognitive evaluation of the agent's trustworthiness. Users make inferences on the agent's trustworthiness by reflecting on issues such as the relevancy of questions raised by the PRA, amount and scope of explanatory information it provided, or how well the recommended products conform to the preference structure users have specified. As an example, users' trusting beliefs in PRAs could be enhanced when PRAs provide additional information in the form of explanations to reveal their underlying reasoning process and cognitively justify their recommendations [125].

In addition to the product-related information and explanations, other sources are also used by consumers to form their trusting beliefs in PRAs, especially by those who have limited product knowledge and therefore cannot accurately appraise the completeness and integrity of the information provided by the agent. When people interact with others for the first time, the initial phase of trust building is based on "whatever information is available" [80]. This information often comes in the form of peripheral cues, such as the trustee's physical appearance [107] or the online vendor's guarantee policies [98]. As Fogg and Tseng [35] have pointed out, the impressions formed through simple inspection of "surface attributes" can directly influence users'

attitudes and behavior. A similar pattern has been found in human–agent interactions: if an embodied agent uses linguistic cues that are widely employed by human beings for building up interpersonal trust, it is perceived as more benevolent and credible [21].

In business-to-consumer (B2C) e-services, social presence is argued as both an enabler and an antecedent of trust, inasmuch as untrustworthy behavior is more easily hidden in lean social presence situations [38]. The significant mediating effect of social presence on Web site content credibility has also been identified by Lee and Nass [65] in a study of synthetic voices in an e-commerce context. Moreover, shoppers' trusting beliefs toward a virtual agent are found to be significantly affected by the perceived social presence of the agent: agents with more lively personifications are perceived more convincing than those with a "cold" visual appearance [47].

In face-to-face human interactions, trustworthiness is usually manifested by a plethora of trivial but important social cues. For example, perceived sociability is considered as one of the dimensions of source credibility [79]. When inexperienced users interact with an agent for the first time, they have very few cues to judge an agent's trustworthiness. A socially rich experience can therefore permit a recommendation agent to function as an information source that not only gives personalized advice but also exhibits a human touch, which is considered more credible than other impersonalized sources [102]. In summary, we propose that

> *Hypothesis 2: Users' perceptions of social presence will positively affect their trusting beliefs.*

Social Presence and Perceived Enjoyment

In traditional retailing environments, salesperson characteristics are found to influence customers' emotions during shopping via buyer–seller interactions [67]. PRA users' perception of enjoyment is an important component of the social benefits they acquire from communicating with the agent. Perceived enjoyment refers to the extent to which the activity of interacting with the recommendation agent is perceived to be enjoyable in its own right aside from the utilitarian value of the agent [29]. It is employed here to assess the hedonic quality of consumer's social interactions with the PRA.

Users' affective responses are found being influenced by the social interface of a software agent [116]. Earlier works have shown that the computer interfaces can attract more attention when human-like characters are employed [115, 121]. Particularly in the domains of computer games and multimedia tutoring systems, the accompanying of an anthropomorphic agent significantly increases the entertainment value of an interface [30]. Such effects could be largely explained by the "fantasy" factor proposed by Malone [75]: a system with fantasy elements can evoke mental images of social situations that are not actually present. By using metaphors of physical objects or situations with which the users are already familiar, designers could create imaginary figures to satisfy users' emotional needs and make users' experiences more enjoyable.

As a major portion of people's lives concerns dealing with other human beings, in general, people like to learn and conduct business with other people present. This is

not only convenient and often necessary, it is also intrinsically desired [10]. Researchers studying physical presence have long found that technologies that afford a strong sense of social presence can provide enjoyment and delight to their users [46, 74]. A significant positive impact of social presence on enjoyment has also been identified in online shopping environments [44]. Therefore, we hypothesize that in human–PRA interactions, people will enjoy interacting with a PRA that carries more social cues and social intelligence.

Hypothesis 3: Users' perceptions of social presence will positively affect their perceived enjoyment.

Antecedents of Agent Usage Intentions

The connections between trust and TAM have been widely discussed in previous studies [39, 72, 97, 124, 126]. Gefen et al. [39] have integrated consumer trust into the traditional TAM model in the context of online shopping, arguing that trust, conceptualized as a set of trusting beliefs toward an online vendor, would directly affect intentions to use a B2C Web site, along with perceived usefulness (PU). Trust in a merchant can also affect PU in both the short term and the long term. Wang and Benbasat [124] have extended this integrated trust–TAM model to the context of online recommendation agents and found that initial trust in online recommendation agents can positively affect usage intentions as well as the consumers' perceived usefulness of the agent. Thus, such relationship is reexamined in this study:[3]

Hypothesis 4: Users' trusting beliefs in a PRA will positively affect their intentions to use the agent.

Hypothesis 5: Users' trusting beliefs in a PRA will positively affect perceived usefulness of the agent.

Hypothesis 6: Perceived usefulness of a PRA will positively affect consumers' intentions to use the agent.

Marketing research has suggested that individuals' shopping behavior is driven by both instrumental and hedonic motivations [4]. In online shopping studies, the effects of shopping enjoyment have been identified as important determinants of online customer loyalty [53, 59]; furthermore, the immersive, hedonic aspects of online shopping environments have been found to play at least an equal role as the instrumental aspects of the environments [23].

Meanwhile, perceived enjoyment is considered to be an important addition to the original TAM [29]. The influence of enjoyment on users' technology adoption intentions have been empirically examined in contexts such as general computer usage [51], Internet usage [81, 117, 119], instant messaging tools [71], Internet-based learning medium [66, 128], and online shopping [59, 120]. Recent research has suggested that perceived enjoyment is not only an immediate antecedent of perceived usefulness, but it is also an important factor influencing the adoption of IS, especially for hedonic IS [120].

The current study proposes that, for users of a recommendation agent, perceived enjoyment will exert an *indirect* influence on their adoption intentions. As argued by Van der Heijden [120], the unmediated effects of perceived enjoyment on adoption intentions are mostly manifested only among hedonic-oriented systems, in which the instrumental values of the system are less important than the pleasurable experience a user could acquire from the usage. Because the primary role of a PRA is to help shoppers complete the cognitive task of identifying a specific product that best meets one's needs from among hundreds of alternatives, extrinsic motivation is still the most important driver of using an agent. When the agent is fun to use, users are likely to be more engaged in the task and therefore more likely to get higher quality advice. Therefore, the effects of perceived enjoyment on behavioral intentions will be fully mediated by perceived usefulness.

Hypothesis 7: Users' perceptions of enjoyment will positively affect perceived usefulness of the agent.

Research Method

Experimental Design

A LABORATORY EXPERIMENT WAS DESIGNED TO EXAMINE the front end of our research model, that is, the direct effects of animated avatar and voice output on social presence. A 2×3 full factorial design was employed using all six combinations of the two independent variables: (1) the presence or absence of an animated face and (2) the output modality of the agent (text, TTS voice, or human voice). To implement the treatments, we designed six interfaces for the same product recommendation agent.

The Agent and Interface Design

The PRA designed for the experiment was a recommendation system for digital cameras, which was adapted and extended from earlier works by Wang and Benbasat [124, 125]. It makes recommendations about digital cameras based on the preferences and needs specified by consumers, applying a content-filtering method. Digital cameras were chosen as the target product due to the complexity of their product attributes, the large number of alternative models available on the market, and the short life span of each generation of products. Previous research has shown that the content-filtering-based agent technology works best for relatively complex products [108]; in real-life practice, most content-filtering PRAs have been designed for attribute-intensive products such as computers, electronics, and automobiles.

The animated characters were designed with Oddcast's SitePal, a Flash-based animation design tool that can integrate both animated characters and speech output with textual content. SitePal provides a wide array of animated characters, and each character can be further customized in terms of its physical appearance features (e.g., face, eyes, nose, hairstyle, makeup, attire, jewelry, and accessories). As one of the

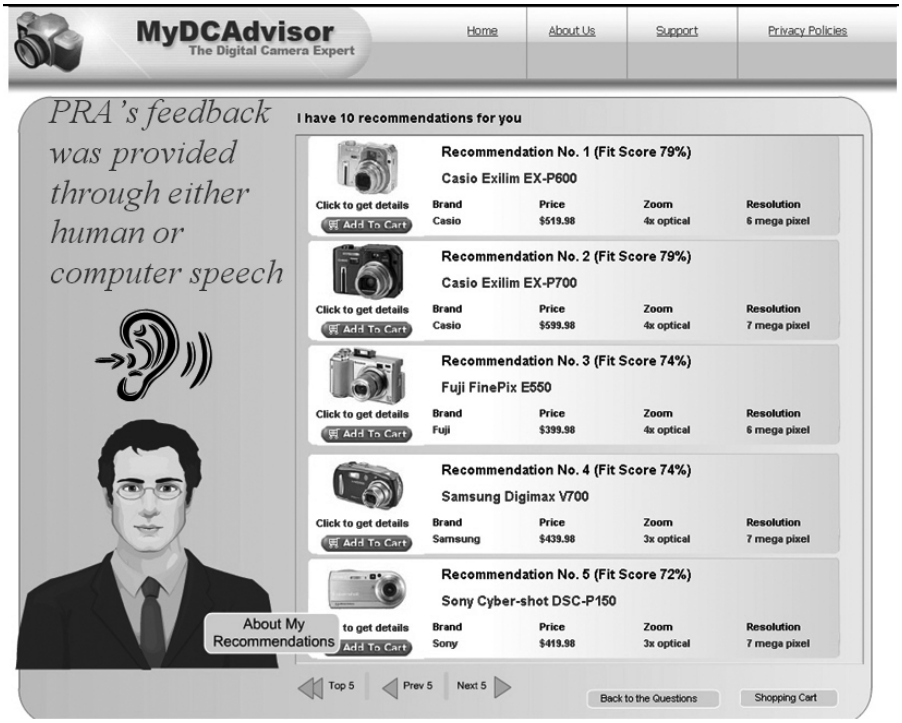

Figure 2. Interface of the With-Face and Speech Output (TTS or Human Voice) Conditions

most widely used plug-ins for Internet browsers, Adobe Flash is an ideal platform for online stores to integrate rich-media elements with HTML-based contents.

Among the three *with-face* conditions, when either TTS voice or human voice was used, the screen layout of the agent was designed to simulate commercial Web sites such as discoveryourride.com (see Figure 2 for the screenshot). The screen layouts of the text-only conditions are shown in Figure 3.

For the three *without-face* conditions, the screen layouts are shown in Figure 4 (with either TTS or human speech output) and Figure 5 (with textual output).

As the particular implementation of the humanoid embodiment may directly affect users' perceptions of the interface, this study planned to examine the effects of the PRA's anthropomorphic interface when a good design is implemented and potential confounds are sufficiently controlled. Our design was guided by the similarity-attraction hypothesis [2], which postulates that people are attracted to others who are similar to them in terms of personality or behavior. For example, Byrne et al. [20] have provided evidence that attraction toward another individual is a positive linear function of the proportion of similar characteristics. Our preliminary analysis also indicated that there were benefits, but no downside, in assigning participants to treatments based on similarity [103]. Therefore, we assigned each participant a PRA with matched-gender (in all conditions) and with matched-ethnicity (in the three with-avatar conditions).

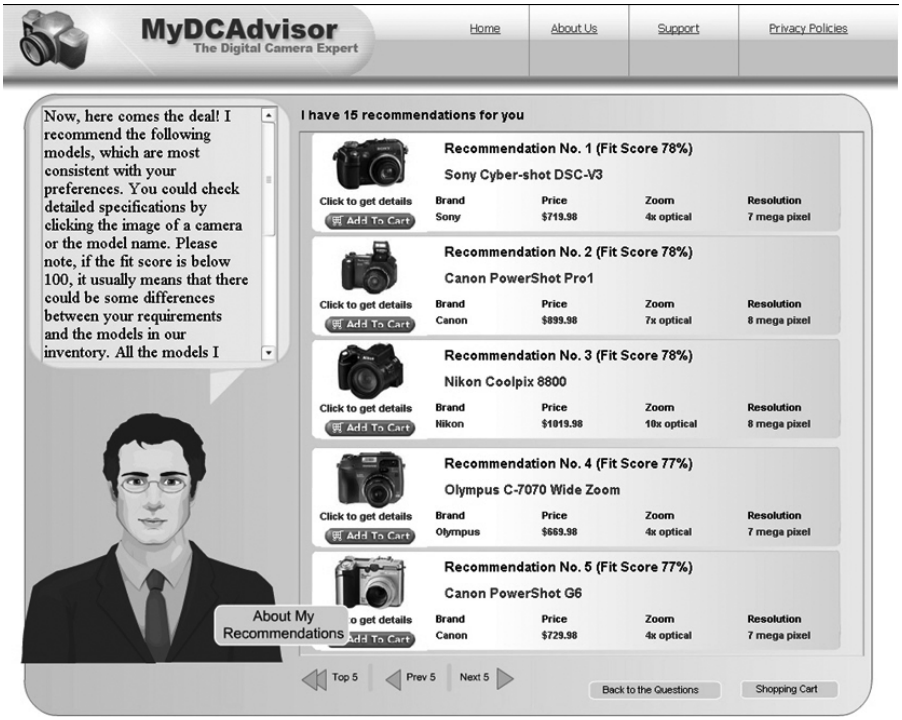

Figure 3. Interface of the With-Face and Textual Output Condition

In the "no avatar/text-only" condition, PRA's gender was controlled by its on-screen name ("Mike" versus "Mary").

In order to select four animated characters exhibiting most salient gender and ethnic characteristics, two rounds of pretests were conducted. In the first round, a sample of 25 graduate students and faculty members were asked to evaluate pictures of 20 candidate characters on their ethnic and gender characteristics. Based on their feedback, four characters (Asian Male, Asian Female, Caucasian Male, and Caucasian Female) were picked. In the second round of pretests, 70 undergraduate and graduate students were asked to evaluate the ethnicity, gender, physical attractiveness, and professionalism of these four characters. The results showed that all of the participants could correctly identify the characters' ethnicity and gender, and that there is no significant difference in the agents' physical attractiveness ($F(3, 207) = 1.044$, $p > 0.30$) and professionalism ($F(3, 207) = 1.560$, $p > 0.20$).

In all six treatment conditions, the participant was asked to select answers to a series of ten questions each associated with an attribute of the camera and to indicate the importance level of this attribute in his or her decision to select a camera. If the participant wanted to know more about this attribute, he or she could click the "About This Question" or "About My Recommendations" button beside the agent to obtain additional explanatory information. To make the human–PRA interaction feel more like a conversation, whenever a user selected an answer, the agent would provide some

Figure 4. Interface of Without-Face and With-Speech (TTS or Human Voice) Conditions

feedback comments, such as acknowledging the user's input or providing additional decisional guidance. After the user completed all ten questions, the agent would offer a list of recommendations sorted by the matching score.

Measures

This study used validated scales for all dependent variables, with minor changes on wording (all measurement items are listed in Appendix B). The measures for social presence were adapted from Gefen et al. [39] and Kumar and Benbasat [60]. The trusting beliefs in recommendation agents were adapted from Wang and Benbasat [125], which in turn relied on Web trust measures developed by McKnight et al. [78]. The measures for perceived enjoyment were adapted from Koufaris [59] and Van der Heijden [120], both of which have been used to measure the extent to which users feel a sense of fun from interacting with an online shopping Web site. In the present study, the "shopping Web site" was replaced by the "recommendation agent." The measures for perceived usefulness and usage intentions were adapted from Davis [28]. Because the constructs were measured by multiple items, summated scales based on the average scores of the multi-items were used in the analysis. Responses were recorded on a seven-point Likert scale (1 = "strongly disagree"; 7 = "strongly agree").

Figure 5. Interface of the Without-Face and Textual-Output Condition

Participants, Incentives, and Experimental Procedure

A total of 168 participants participated in this study, which corresponds to 28 subjects per group. According to Cohen [25], the choice of the sample size is to ensure sufficient statistical power (about 80 percent) at the significance level of 0.05 when medium effect size ($f = 0.25$) was assumed for the main and interaction effects. Participants are students or staff members recruited through online and offline advertisements at a large North American university. To control the gender and ethnicity factors, we invited equal numbers of participants from four demographic groups: female Asian, female Caucasian, male Asian, and male Caucasian. They were then proportionally assigned to one of the six treatment groups and with a gender- and ethnicity-matched agent. To avoid potential biases in their evaluations and to make interactions with the PRA more necessary to complete the task, only individuals who did not own or had not purchased a digital camera before, as well as those who had owned or purchased a digital camera but considered themselves to have low expertise, were recruited [124].[4]

The experiment was administered during a one-hour session for each participant. Each participant was welcomed and given an "informed consent" form to review. He or she was then asked to fill out a short online questionnaire which included demographic questions and instruments to measure all control variables. A research assistant then trained participants on how to use the agent by playing a video clip with narration that

explained the interface and all functionalities of the PRA. Afterward, the participant was asked to complete a shopping task. To reduce the potential confound that the subject may have purchased a digital camera for a similar situation before, two shopping tasks were prepared and the participant was randomly assigned to one of them. One task asked the subject to choose a digital camera for a photography enthusiast; the other asked the subject to select a model for an amateur user (post hoc tests indicated no significant differences between tasks on all dependent variables). After the task, participants were asked to complete a posttask questionnaire which included measures of all dependent variables.

Each participant was guaranteed a monetary compensation for his or her participation ($10). In order to motivate participants to view the experiment as a real-life shopping task and to increase their involvement, they were told that their choices would be ranked by a digital camera expert on its appropriateness for the requirements described in the task scenarios and the top 25 percent best performers could receive an extra $50 bonus.

Data Analysis

Of the 168 participants, there were 155 undergraduate students, 8 graduate students, and 5 staff members. One hundred sixty-six of them had more than four years of Internet experience, within which 135 had used the Internet for more than seven years.

In the background questionnaire, participants were asked about their product expertise and perceived risk of online shopping. Their comfort levels with the Internet and online shopping were also assessed. No significant differences were found among subjects randomly assigned to the six experimental groups for each of these variables. In addition, no significant differences were found among groups with respect to participants' age and years of Internet experience.

To perform our analysis, we used univariate analysis of variance (ANOVA) to examine the direct effect of animated avatar and voice output on social presence. For the causal model, we used the variance-based partial least squares (PLS) tool so as to simultaneously examine the structural component and the measurement component in one model. Construct reliability was examined by Cronbach's alpha and composite reliability. All constructs met the benchmark of acceptable reliability ($\alpha > 0.70$) [95]. Convergent validity was examined by the average variance extracted (AVE) for each construct. Fornell and Larcker [36] suggested that the AVE from a construct should exceed 0.5. As shown in Table 1, all constructs satisfied this criterion. Discriminant validity was examined by two criteria. One criterion is that a construct should share more variance with its own measures than it shares with other constructs in a model. As shown in Table 1, the square roots of the AVE were greater than the correlations between the construct with other constructs [6]. The second criterion is that no item should load more highly with a construct other than the construct it intends to measure. The loadings and cross-loadings of measures are shown in Table 2. An examination of the matrix revealed that all items satisfied this criterion.

Table 1. Construct Attributes

	Cronbach's alpha	Composite reliability	1	2	3	4	5
1. Social presence	0.887	0.917	0.830				
2. Trusting beliefs	0.866	0.894	0.384	0.663			
3. Perceived enjoyment	0.872	0.909	0.630	0.367	0.817		
4. Perceived usefulness	0.856	0.903	0.386	0.549	0.531	0.837	
5. Usage intentions	0.951	0.969	0.375	0.525	0.434	0.728	0.955

* The scores in the diagonal are square roots of AVEs; the lower triangle represents the correlations between constructs.

Table 2. Loadings and Cross-Loadings of Measures

	Social presence	Trusting beliefs	Perceived enjoyment	Perceived usefulness	Agent usage intentions
SP1	0.854	0.352	0.500	0.363	0.329
SP2	0.799	0.316	0.519	0.351	0.396
SP3	0.872	0.319	0.584	0.335	0.321
SP4	0.791	0.315	0.501	0.292	0.252
SP5	0.830	0.300	0.522	0.278	0.259
TRUST1	0.230	0.600	0.218	0.449	0.404
TRUST2	0.202	0.570	0.224	0.396	0.363
TRUST3	0.175	0.710	0.170	0.413	0.416
TRUST4	0.210	0.625	0.200	0.308	0.342
TRUST5	0.304	0.771	0.247	0.449	0.433
TRUST6	0.353	0.671	0.328	0.393	0.278
TRUST7	0.397	0.709	0.243	0.335	0.298
TRUST8	0.272	0.725	0.311	0.440	0.314
TRUST9	0.308	0.673	0.302	0.352	0.292
TRUST10	0.190	0.694	0.230	0.285	0.333
PE1	0.520	0.311	0.871	0.431	0.356
PE2	0.552	0.280	0.741	0.442	0.399
PE3	0.507	0.314	0.781	0.430	0.316
PE4	0.411	0.317	0.796	0.471	0.375
PE5	0.532	0.274	0.887	0.419	0.324
PU1	0.384	0.470	0.539	0.886	0.678
PU2	0.194	0.437	0.327	0.761	0.529
PU3	0.353	0.460	0.414	0.881	0.630
PU4	0.354	0.527	0.494	0.834	0.595
AUI1	0.381	0.525	0.438	0.712	0.958
AUI2	0.325	0.511	0.385	0.690	0.942
AUI3	0.369	0.474	0.421	0.689	0.965

ANOVA Results for Social Presence

The ANOVA results are summarized in Table 3. Users reported a stronger sense of *social presence* when the agent was embodied with humanoid embodiment ($F(1, 162) = 5.871$, $p < 0.05$). Output modality also exerted a significant influence on social presence ($F(2, 162) = 11.495$, $p < 0.01$), but there was no interaction between humanoid embodiment and output modality ($F(1, 162) = 1.016$, $p > 0.05$). To compare the three output modalities, a Scheffé test for multiple comparisons was conducted (see Table 4), the results of which indicated that the difference between human voice and text, as well as the difference between human voice and TTS voice, are significant, while there is no significant difference between text and TTS voice. Therefore, H1a, H1b, and H1c are supported, but H1d and H1e are not.

Table 3. ANOVA Results on Social Presence

	Degrees of freedom	*F*	*p*-value
Humanoid embodiment	1	5.871	0.016
Output modality	2	11.495	< 0.001
Embodiment × modality	2	1.016	0.364

Table 4. Scheffé Test Results on Output Modality

A	B	Mean difference	*p*-value
Text only	TTS voice	0.243	0.453
	Human voice	–0.650	0.004
TTS voice	Text only	–0.243	0.453
	Human voice	–0.893	< 0.001

PLS Analysis

PLS was used to investigate the causal relationship between social presence and users' intentions to adopt the PRA as a decision aid. As suggested by Chin [24], bootstrapping was performed to test the statistical significance of each path coefficient using *t*-tests. As shown in Figure 6, the hypothesized paths from social presence to trusting beliefs and perceived enjoyment are significant. Besides, the path coefficients between trust, perceived usefulness, perceived enjoyment, and agent usage intentions are all significant. Approximately 55 percent of the variance in the usage intentions is accounted for by the variables in the model ($R^2 = 0.547$). The model was also examined with mediation tests suggested by Baron and Kenny [8], the results of which are shown in Appendix C. Thus, H2, H3, H4, H5, H6, and H7 are supported.

Discussion and Future Research

Discussions of Results

This study found strong evidence for the influence of humanoid embodiments and output modalities on enhancing social interactions between a PRA and its users. As expected, PRAs with animated human faces invoked significantly stronger perceptions of social presence than disembodied ones. In addition, PRAs with voice output induced stronger social responses from users due to the rich social cues embedded in human voice. In contrast, TTS voice, which could only carry limited social cues, was perceived no different from text. All of these findings supported our expectations that humanoid embodiments and voice output could be effectively applied in an online shopping environment and, more particularly, for the interface of advice-giving systems.

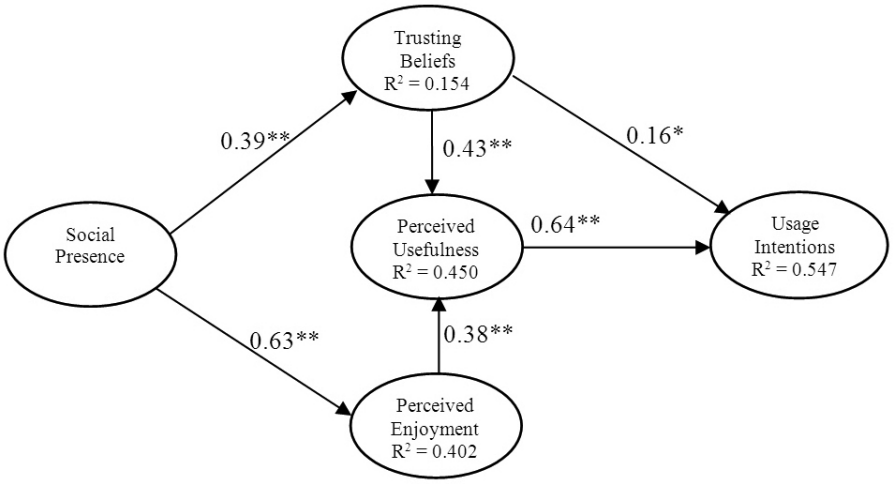

Figure 6. PLS Structural Model (n = 168)

It is also noteworthy that, contrary to our expectations, users in TTS conditions reported lower (though *not* statistically significant) social presence than those in text-only conditions ($Mean_{TTS}$ = 3.79, $Mean_{Text}$ = 4.04, seven-point scale), which suggested that the mere addition of voice channel to the agent interface is not sufficient for boosting social interactions between the user and the PRA. Even though people may be able to attribute certain social characteristics to computer-generated voices, as demonstrated in prior studies [61, 62, 86, 87], it is still premature to apply TTS voice in recommendation agent interface where a large amount of information needs to be delivered. In this case, users are not only repeatedly reminded of the voice's mechanical nature during the lengthy speech playback, they also tend to ignore the voice's social implications when comprehensibility is at threat.

This study further justified the importance of interface design in users' adoption of online recommendation agents through the integration of social presence, trust, and perceived enjoyment along with traditional TAM constructs. Whereas the effects of social presence have been explored in prior studies of e-mail [37, 114] and online shopping Web sites [27, 38], it was in this study extended to the realm of online recommendation systems. Besides, our findings provided more empirical support for the interrelationships among trust, perceived enjoyment, and TAM constructs. As important, this study demonstrated not only that creating perceptions of social presence is important in the context of PRA use, but that such perceptions can be enhanced by carefully designing IT artifacts [14] based on the principles of human–computer interaction as discussed in the second section.

In regard to its practical implications, this study showed that effective social interactions can be implemented with relatively low cost.[5] For online vendors, a 2D Flash-based animated face is probably good enough to embody the types of PRAs studied in this paper. As illustrated by our experiment, Adobe Flash could be an effective platform of anthropomorphic interface design. Considering the very large size of Adobe Flash's

installation base,[6] we believe our findings could effectively relieve online retailers' concerns that their customers may be unable to see the anthropomorphic agent on their browsers without installing additional software. The implication for PRA designers is that social presence could be enhanced by implementing a reasonable degree of anthropomorphism. On the other hand, while it is evident that human speech is more effective than either on-screen text or TTS voice technology, its inflexibility (given that all audio contents need to be prerecorded) and significantly higher development cost[7] may make it infeasible for small or medium-sized online retailers, especially when their PRAs are expected to produce a large volume of interactive content. Meanwhile, the relatively cheap TTS voice[8] was perceived to be no better than text. Therefore, until a more naturally sounding TTS engine is ready for the mainstream market, on-screen text is probably a safer bet for online vendors, especially when content flexibility is imperative or prerecorded human voices are not affordable. Besides, as we did not find that the consistency issue—that is, the interaction between modality of output and human embodiment—was significant, it is reasonable to assume that the effects of these two interface factors are fairly linear and complementary. For online vendors under budget constraints, they could begin with a relatively simple design with partial anthropomorphic interface applied, such as a PRA with humanoid embodiment plus textual output.

Limitations

There are a number of limitations to this study that should be noted. First, our studies have only examined users' first impressions of various implementations of embodied PRAs; however, the contribution of social presence to the building of an enduring consumer–PRA relationship could be moderated by the factor of timing. As studies examining the social effects of computer-mediated communications have revealed, the social attractions induced by seeing the photos of the communication partners are only significant between new and unacquainted partners but not significant between long-term team members [123]. Therefore, the perceptions of social presence may vary as an online consumer becomes more familiar with a PRA. First-time and repeated customers may exhibit different preferences for the design elements that can induce social responses toward a PRA. Therefore, future research should employ a longitudinal design to determine whether the finding still holds when novelty wears off.

The controlled laboratory experiments used in this study, with their advantages on ensuring internal validity, may affect the generalizablity of the results. Given the artificial settings of the fictitious shopping tasks and the simulated PRA, the decision-making processes employed by the subjects during the experiment may differ from if they were actually making a real purchase decision. However, our incentive mechanism (particularly the possibility of a $50 bonus based on the appropriateness of the selected product) helped increase the task relevance as well as the task involvement. Last, because the participants were mostly university students (whose demographic data generally match with those of average Internet users [54]), the results from our

studies may be somewhat different from results obtained using people from other demographic groups. Also, only one type of product was used. Further research with different participant samples and different categories of products is warranted.

Future Research

As users' interactions with PRAs consist of behavioral and psychological activities involving cognitive, social, relational, and affective aspects, further research should study how these activities may interrelate with each other. For example, researchers could explore whether or not positive social and affective responses could compensate for the deficiencies on cognitive processing, and in particular, to what extent an attractive interface agent could help alleviating users' concerns if the recommended products are not convincingly justified or the explanatory information provided is insufficient. Considering the comparatively lower cost of providing textual information, should online stores employ an information-rich agent or a media-rich one? How may customers' motivations and capabilities [100] influence their preferences for a particular PRA design?

In addition, even though we did not identify significant social benefits of using TTS voice on a PRA interface, it is still worthwhile to further investigate other scenarios where TTS may be applicable for social interaction design, in particular those situations where limited amount of information needs to be transmitted; for example, an agent simply designed as a virtual host greeting visitors or announcing promotional offerings. Research is also needed to examine input modalities on the user side. For example, researchers could examine how users' perceptions of social interactions may get enhanced if they could use natural language (either by a keyboard or through speech recognition software) to interact with a recommendation agent, as compared to the current input method of using mouse-click to pick answers for the predefined questions. Last, it will be interesting to examine the applicability of humanoid embodiment and voice output on nonshopping Web sites, such as the ones dedicated to online health interventions and health communication campaigns.

In conclusion, social interactions with online recommendation agents have been recognized as an important issue in online shopping environments. As users' social, relational, and affective needs are increasingly found to directly relate to their assessment of a technology's usefulness and ultimately their intentions to adopt the technology, an agent will be more likely to be accepted and employed by online shoppers when it affords a compelling socially rich and enjoyable interaction experience, on top of its underlying functionalities of product search and comparisons. Additional studies in this area will not only help researchers extend their understandings on how people interact with and respond to emerging technologies in everyday life, but also benefit online retailers by creating a more satisfactory experience for their customers.

Acknowledgments: The authors thank the Natural Science and Engineering Research Council of Canada (NSERC) and the Social Science and Humanities Research Council of Canada (SSHRC) for their support of this research.

Notes

1. Anthropomorphizing an agent entails adding humanlike characteristics, such as facial expressions, speech output, body gestures, auditory and kinesthetic feedback, human emotions, and social intelligence (i.e., emulations of common social conventions). Such humanoid interface characters are also known as avatars.

2. We thank one of the reviewers for pointing this out.

3. The effect of perceived ease of use (PEOU) on behavioral intentions was not examined in this research not only because it has been investigated in numerous TAM-related studies but also due to the focus of this paper. We do not intend to reexamine the interrelationships between all TAM-related constructs; instead, we only focus on those constructs that may be directly or indirectly affected by the two interface components via influencing social presence. We had not hypothesized any direct relationship between our interface components and PEOU because we believe that various implementations of the PRA in our study would not create different levels of cognitive load as the agents have asked identical product-related questions and used the same interface for answering questions, the same output format, the same input/output functionality used by users, etc., as well as utilized the same underlying algorithm. Besides, as no existing theories or empirical findings have projected the direct relationship between social presence and PEOU, we have not hypothesized such relationship either.

4. Subjects were asked to report their expertise level on digital cameras when signing up for the study. A seven-point scale was used, ranging from 1 = "not at all" to 7 = "expert." Only those with a score of 4 or below were invited.

5. Adding a Flash-based virtual character to a Web site could be as low as $10.00 per month by using third-party application service providers such as Oddcast Inc. More charges will be applied when audio capabilities are enabled.

6. The penetration rate of Adobe Flash Player is over 98 percent, according to a recent survey conducted by Adobe (www.adobe.com/products/player_census/flashplayer/PC.html).

7. Hiring a professional narrator for human speech recording could cost at least $50 per minute of speech (price quoted by Oddcast Inc.), and it usually takes 2–3 business days if not done in-house.

8. For example, $100 extra per year for TTS functionality and $4 for 1,000 audio streams (price quoted by Oddcast Inc.).

References

1. Alba, J.; Lynch, J.; Weitz, B.; Janiszewski, C.; Lutz, R.; Sawyer, A.; and Wood, S. Interactive home shopping: Consumer, retailer, and manufacturer incentives to participate in electronic marketplaces. *Journal of Marketing, 61,* 3 (1997), 38–53.

2. Al-Natour, S.; Benbasat, I.; and Cenfetelli, R. The role of design characteristics in shaping perceptions of similarity: The case of online shopping assistants. *Journal of the Association for Information Systems, 7,* 12 (2006), 821–861.

3. Argyle, M. *Bodily Communication,* 2d ed. New York: International Universities Press, 1988.

4. Babin, B.J.; Darden, W.R.; and Griffen, M. Work and/or fun: Measuring hedonic and utilitarian shopping value. *Journal of Consumer Research, 20,* 4 (1994), 644–656.

5. Bailenson, J.N.; Swinth, K.; Hoyt, C.; Persky, S.; Dimov, A.; and Blascovich, J. 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: Teleoperators and Virtual Environments, 14,* 4 (2005), 379–393.

6. Barclay, D.; Thompson, R.; and Higgins, C. The partial least squares (PLS) approach to causal modeling: Personal computer adoption and use as an illustration. *Technology Studies, 2,* 2 (1995), 285–309.

7. Barlow, A.K.J.; Siddiqui, N.Q.; and Mannion, M. Development in information and communication technologies for retail marketing channels. *International Journal of Retail and Distribution Management, 32,* 2 (2004), 157–163.

8. Baron, R.M., and Kenny, D.A. The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51,* 6 (1986), 1173–1182.

9. Bartneck, C., and Reichenbach, J. Subtle emotional expressions of synthetic characters. *International Journal of Human–Computer Studies, 62,* 2 (2005), 179–192.

10. Baumeister, R.F., and Leary, M.R. The need to belong: Desire for interpersonal attachments as a fundamental human motivation. *Psychology Bulletin, 117,* 3 (1995), 497–529.

11. Baylor, A.L. The impact of pedagogical agent image on affective outcomes. Paper presented at the International Conference on Intelligent User Interface, Workshop on Affective Interactions: Computers in the Affective Loop, San Diego, CA, 2005.

12. Baylor, A.L., and Ryu, J. The effects of image and animation in enhancing pedagogical agent persona. *Journal of Educational Computing Research, 28,* 4 (2003), 373–394.

13. Benbasat, I., and Zmud, R.W. Empirical research in information systems: The practice of relevance. *MIS Quarterly, 23,* 1 (1999), 3–16.

14. Benbasat, I., and Zmud, R.W. The identity crisis within the IS discipline: Defining and communicating the discipline's core properties. *MIS Quarterly, 27,* 2 (2003), 183–194.

15. Berry, D.C.; Butler, L.T.; and de Rosis, F. Evaluating a realistic agent in an advice-giving task. *International Journal of Human–Computer Studies, 63,* 3 (2005), 304–327.

16. Bickmore, T., and Cassell, J. Relational agents: A model and implementation of building user trust. In J.A. Jacko, A. Sears, M. Beaudouin-Lafon, and R.J.K. Jacob (eds.), *Proceedings of the 2001 SIGCHI Conference on Human Factors in Computing Systems.* New York: ACM Press, 2001, pp. 396–403.

17. Biocca, F. The cyborg's dilemma: Progressive embodiment in virtual environments. *Journal of Computer Mediated Communication, 3,* 2 (1997) (available at http://jcmc.indiana.edu/vol3/issue2/biocca2.html).

18. Biocca, F.; Harms, C.; and Burgoon, J.K. Towards a more robust theory and measure of social presence: review and suggested criteria. *Presence: Teleoperators and Virtual Environments, 12,* 5 (2003), 456–480.

19. Burke, R.R. Technology and the customer interface: What consumers want in the physical and virtual store. *Journal of the Academy of Marketing Science, 31,* 2 (2002), 109–126.

20. Byrne, D.; Griffitt, W.; and Stefaniak, D. Attraction and similarity of personality characteristics. *Journal of Personality and Social Psychology, 5,* 1 (1967), 82–90.

21. Cassell, J., and Bickmore, T. External manifestations of trustworthiness in the interface. *Communications of the ACM, 43,* 12 (2000), 50–56.

22. Cassell, J.; Sullivan, J.; Prevost, S.; and Churchill, E. *Embodied Conversational Agents.* Cambridge, MA: MIT Press, 2000.

23. Childers, T.L.; Carr, C.L.; Peck, J.; and Carson, S. Hedonic and utilitarian motivations for online retail shopping behavior. *Journal of Retailing, 77,* 4 (2001), 511–535.

24. Chin, W. The partial least squares approach to structural equation modeling. In G.A. Marcoulides (ed.), *Modern Methods for Business Research.* Hillsdale, NJ: Lawrence Erlbaum, 1998, pp. 295–336.

25. Cohen, J. *Statistical Power Analysis for the Behavioral Sciences,* 2d ed. Hillsdale, NJ: Lawrence Erlbaum, 1988.

26. Cowell, A.J., and Stanney, K.M. Manipulation of non-verbal interaction style and demographic embodiment to increase anthropomorphic computer character credibility. *International Journal of Human–Computer Studies, 62,* 2 (2005), 281–306.

27. Cyr, D.; Hassanein, K.; Head, M.; and Ivanov, A. The role of social presence in establishing loyalty in e-service environments. *Interacting with Computers, 19,* 1 (2007), 43–56.

28. Davis, F.D. Perceived usefulness, perceived ease of use and user acceptance of information technology. *MIS Quarterly, 13,* 3 (1989), 319–339.

29. Davis, F.D.; Bagozzi, R.P.; and Warshaw, P.R. Extrinsic and intrinsic motivation to use computers in the workplace. *Journal of Applied Social Psychology, 22,* 14 (1992), 1111–1132.

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

31. Dempsey, J.V., and Van Eck, R. Modality and placement of a pedagogical adviser in individual interactive learning. *British Journal of Educational Technology, 34,* 5 (2003), 1–16.

32. de Rosis, F.; Pelachaud, C.; Poggi, I.; Carofiglio, V.; and De Carolis, B. From Greta's mind to her face: Modeling the dynamics of affective states in a conversational embodied agent. *International Journal of Human–Computer Studies, 59,* 1–2 (2003), 81–118.

33. Dutoit, T. *An Introduction to Text-To-Speech Synthesis (Text, Speech and Language Technology, V. 3).* Dordrecht: Kluwer Academic, 1997.

34. Fiske, S.T., and Taylor, S.E. *Social Cognition.* New York: McGraw-Hill, 1991.

35. Fogg, B.J., and Tseng, H. The elements of computer credibility. In *Proceedings of the 1999 SIGCHI Conference on Human Factors in Computing Systems.* New York: ACM Press, 1999, pp. 80–87.

36. Fornell, C., and Larcker, D. Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research, 18,* 1 (1981), 39–50.

37. Gefen, D., and Straub, D. Gender differences in perception and adoption of e-mail: An extension to the technology acceptance model. *MIS Quarterly, 21,* 4 (1997), 389–400.

38. Gefen, D., and Straub, D. Managing user trust in B2C e-services. *e-Service Journal, 2,* 2 (2003), 7–24.

39. Gefen, D.; Karahanna, E.; and Straub, D.W. Trust and TAM in online shopping: An integrated model. *MIS Quarterly, 27,* 1 (2003), 51–90.

40. Giles, D.C. Parasocial interaction: A review of the literature and a model for future research. *Media Psychology, 4,* 3 (2002), 279–305.

41. Giles, H. Evaluative reactions to accents. *Educational Review, 22,* 3 (1970), 211–227.

42. Giles, H., and Powesland, P.F. *Speech Style and Social Evaluation.* New York: Academic Press, 1975.

43. Grant, A.E.; Guthrie, K.K.; and Ball-Rokeach, S.J. Television shopping: A media system dependency perspective. *Communication Research, 18,* 6 (1991), 773–798.

44. Hassanein, K., and Head, M. The impact of infusing social presence in the Web interface: An investigation across product types. *International Journal of Electronic Commerce, 10,* 2 (2006), 31–55.

45. Haubl, G., and Trifts, V. Consumer decision making in online shopping environments: The effects of interactive decision aids. *Marketing Science, 19,* 1 (2000), 4–21.

46. Heeter, C. Communication research on consumer VR. In F. Biocca and M.R. Levy (eds.), *Communication in the Age of Virtual Reality.* Hillsdale, NJ: Lawrence Erlbaum, 1995, pp. 191–218.

47. Hertzum, M.; Andersen, H.H.K.; Andersen, V.; and Hansen, C.B. Trust in information sources: Seeking information from people, documents, and virtual agents. *Interacting with Computers, 14,* 5 (2002), 575–599.

48. Hevner, A.R.; March, S.T.; Park, J.; and Ram, S. Design science in information systems research. *MIS Quarterly, 28,* 1 (2004), 77–105.

49. Holzwarth, M.; Janiszewski, C.; and Neumann, M.M. The influence of avatars on online consumer shopping behavior. *Journal of Marketing, 70,* 10 (2006), 19–36.

50. Ichheiser, G. *Appearance and Realities: Misunderstanding in Human Relations.* San Francisco: Jossey-Bass, 1970.

51. Igbaria, M.; Iivari, J.; and Maragahh, H. Why do individuals use computer technology? A Finnish case study. *Information & Management, 29,* 5 (1995), 227–238.

52. Isbister, K., and Nass, C. Consistency of personality in interactive characters: Verbal cues, non-verbal cues, and user characteristics. *International Journal of Human–Computer Studies, 53,* 2 (2000), 251–267.

53. Jarvenpaa, S.L., and Todd, P.A. Consumer reactions to electronic shopping on the World Wide Web. *International Journal of Electronic Commerce, 1,* 2 (1997), 59–88.

54. Johnson, C.A. The North American consumer: Online retail update. Consulting Report, Forrester Research, Cambridge, MA, 2005.

55. Keeling, K.; McGoldrick, P.; Beatty, S.; and Macaulay, L. Face value? Customer views of appropriate formats for embodied conversational agents (ECAs) in online retailing. In R.H. Sprague Jr. (ed.), *Proceedings of the Thirty-Seventh Annual Hawaii International Conference on System Sciences.* Los Alamitos, CA: IEEE Computer Society, 2004, pp. 178–187.

56. Klein, J.; Moon, Y.; and Picard, R.W. This computer responds to user frustration: Theory, design, and results. *Interacting with Computers, 14,* 2 (2002), 119–140.

57. Koda, T., and Maes, P. Agents with faces: The effects of personification. In R. Bilger, S. Guest, and M.J. Tauber (eds.), *Proceedings of the Fifth IEEE International Workshop on Robot and Human Communication.* Los Alamitos, CA: IEEE Computer Society, 1996, pp. 189–194.

58. Komiak, S.Y.X., and Benbasat, I. The effects of personalization and familiarity on trust and adoption of recommendation agents. *MIS Quarterly, 30,* 4 (2006), 941–960.

59. Koufaris, M. Applying the technology acceptance model and flow theory to online consumer behavior. *Information Systems Research, 13,* 2 (2002), 205–233.

60. Kumar, N., and Benbasat, I. The influence of recommendations and consumer reviews on evaluations of Websites. *Information Systems Research, 17,* 4 (2006), 425–439.

61. Lee, E.-J. Effects of "gender" of the computer on informational social influence: The moderating role of task type. *International Journal of Human-Computer Studies, 58,* 4 (2003), 347–362.

62. Lee, E.-J.; Nass, C.; and Brave, S. Can computer-generated speech have gender? An experimental test of gender stereotype. In *Proceedings of the 2000 ACM SIGCHI Conference on Human Factors in Computing Systems: Extended Abstracts.* New York: ACM Press, 2000, pp. 289–290.

63. Lee, K.M. Why presence occurs: Evolutionary psychology, media equation, and presence. *Presence: Teleoperators and Virtual Environments, 13,* 4 (2004), 494–505.

64. Lee, K.M., and Nass, C. Designing social presence of social actors in human–computer interaction. In *Proceedings of the 2003 SIGCHI Conference on Human Factors in Computing Systems.* New York: ACM Press, 2003, pp. 289–296.

65. Lee, K.M., and Nass, C. The multiple source effect and synthesized speech: Doubly-disembodied language as a conceptual framework. *Human Communication Research, 30,* 2 (2004), 182–207.

66. Lee, M.K.O.; Cheung, C.M.K.; and Chen, Z. Acceptance of Internet-based learning medium: The role of extrinsic and intrinsic motivation. *Information & Management, 42,* 8 (2005), 1095–1104.

67. Lee, S., and Dubinsky, A.J. Influence of salesperson characteristics and customer emotion on retail dyadic relationships. *International Review of Retail, Distribution & Consumer Research, 13,* 1 (2003), 21–36.

68. Lee, W.-P.; Liu, C.-H.; and Lu, C.-C. Intelligent agent-based systems for personalized recommendations in Internet commerce. *Expert Systems with Applications, 22,* 4 (2002), 275–284.

69. Lester, J.C., and Stone, B.A. Increasing believability in animated pedagogical agents. In *Proceedings of the First International Conference on Autonomous Agents.* New York: ACM Press, 1997, pp. 16–21.

70. Lester, J.C.; Converse, S.A.; Kahler, S.E.; Barlow, S.T.; Stone, B.A.; and Bhogal, R.S. The persona effect: Affective impact of animated pedagogical agents. In *Proceedings of the 1997 SIGCHI Conference on Human Factors in Computing Systems.* New York: ACM Press, 1997, pp. 359–366.

71. Li, D.; Chau, P.Y.K.; and Lou, H. Understanding individual adoption of instant messaging: An empirical investigation. *Journal of the Association for Information Systems, 6,* 4 (2005), 102–129.

72. Lim, K.H.; Sia, C.-L.; Lee, M.K.O.; and Benbasat, I. Do I trust you online, and if so, will I buy? An empirical study of two trust-building strategies. *Journal of Management Information Systems, 23,* 2 (Fall 2006), 233–266.

73. Link, K.E.; Kreuz, R.J.; Graesser, A.C.; and Group, T.T.R. Factors that influence the perception of feedback delivered by a pedagogical agent. *International Journal of Speech Technology, 4,* 2 (2001), 145–153.

74. Lombard, M., and Ditton, T. At the heart of it all: The concept of presence. *Journal of Computer Mediated Communication, 3,* 2 (1997) (available at http://jcmc.indiana.edu/vol3/issue2/lombard.html).

75. Malone, T.W. Heuristics for designing enjoyable user interface: Lessons from computer games. In J.C. Thomas and M.L. Schneider (eds.), *Human Factors in Computer Systems.* Norwood, NJ: Ablex, 1984, pp. 1–12.

76. Mayer, R.E.; Sobko, K.; and Mautone, P.D. Social cues in multimedia learning: Role of speaker's voice. *Journal of Educational Psychology, 95,* 2 (2003), 419–425.

77. McBreen, H.M., and Jack, M.A. Evaluating humanoid synthetic agents in e-retail applications. *IEEE Transactions on Systems, Man, and Cybernetics—Part A: Systems and Humans, 31,* 5 (2001), 394–405.

78. McKnight, D.H.; Choudhury, V.; and Kacmar, C. Developing and validating trust measures for e-commerce: An integrative typology. *Information Systems Research, 13,* 3 (2002), 334–359.

79. Metzger, M.J.; Flanagin, A.J.; Eyal, K.; Lemus, D.R.; and Mccann, R.M. Credibility for the 21st century: Integrating perspectives on source, message, and media credibility in the contemporary media environment. In P. Kalbfleisch (ed.), *Communication Yearbook 27.* Mahwah, NJ: Lawrence Erlbaum, 2003, pp. 293–335.

80. Meyerson, D.; Weick, K.E.; and Kramer, R.M. Swift trust and temporary groups. In R.M. Kramer and T.R. Tyler (eds.), *Trust in Organizations: Frontiers of Theory and Research.* Thousand Oaks, CA: Sage, 1996, pp. 261–287.

81. Moon, J.-W., and Kim, Y.-G. Extending the TAM for a World-Wide-Web context. *Information & Management, 38,* 4 (2001), 217–230.

82. Moreno, R.; Mayer, R.E.; Spires, H.A.; and Lester, J.C. The case for social agency in computer-based teaching: Do students learn more deeply when they interact with animated pedagogical agents? *Cognition and Instruction, 19,* 2 (2001), 177–213.

83. Moundridou, M., and Virvou, M. Evaluating the persona effect of an interface agent in a tutoring system. *Journal of Computer Assisted Learning, 18,* 3 (2002), 253–261.

84. Nass, C., and Gong, L. Maximized modality or constrained consistency? Paper presented at the International Conference on Auditory-Visual Speech Processing (AVSP'99), Santa Cruz, CA, August 7–10, 1999.

85. Nass, C., and Gong, L. Speech interfaces from an evolutionary perspective. *Communications of the ACM, 43,* 9 (2000), 36–43.

86. Nass, C., and Lee, K.M. Does computer-generated speech manifest personality? An experimental test of similarity attraction. In *Proceedings of the 2000 SIGCHI Conference on Human Factors in Computing Systems.* New York: ACM Press, 2000, pp. 329–336.

87. Nass, C., and Lee, K.M. Does computer-synthesized speech manifest personality? Experimental tests of recognition, similarity-attraction, and consistency-attraction. *Journal of Experimental Psychology: Applied, 7,* 3 (2001), 171–181.

88. Nass, C., and Moon, Y. Machines and mindlessness: Social responses to computers. *Journal of Social Issues, 56,* 1 (2000), 81–103.

89. Nass, C., and Scott, B. *Wired for Speech: How Voice Activates and Advances the Human–Computer Relationship.* Cambridge, MA: MIT Press, 2005.

90. Nass, C., and Steuer, J. Voices, boxes, and sources of messages: Computers and social actors. *Human Communication Research, 19,* 4 (1993), 504–527.

91. Nass, C.; Isbister, K.; and Lee, E.-J. Truth is beauty: Researching embodied conversational agents. In J. Cassell, J. Sullivan, S. Prevost, and E. Churchill (eds.), *Embodied Conversational Agents.* Cambridge, MA: MIT Press, 2000, pp. 374–402.

92. Nass, C.; Kim, E.-Y.; and Lee, E.-J. When my face is the interface: An experimental comparison of interacting with one's own face or someone else's face. In C.-M. Karat, A. Lund, J. Coutaz, and J. Karat (eds.), *Proceedings of the 1998 SIGCHI Conference on Human Factors in Computing Systems.* New York: ACM Press, 1998, pp. 148–154.

93. Nass, C.; Moon, Y.; and Green, N. Are computers gender-neutral? Gender stereotypic responses to computers. *Journal of Applied Social Psychology, 27,* 10 (1997), 864–876.

94. Nowak, K.L., and Biocca, F. The effect of the agency and anthropomorphism on user's sense of telepresence, copresence, and social presence in virtual environments. *Presence: Teleoperators and Virtual Environments, 12,* 5 (2003), 481–494.

95. Nunnally, J.C. *Psychometric Theory,* 2d ed. New York: McGraw-Hill, 1978.

96. Ochsman, R.B., and Chapanis, A. The effects of 10 communication modes on the behavior of teams during co-operative problem-solving. *International Journal of Man–Machine Studies, 6,* 5 (1974), 579–619.

97. Pavlou, P.A. Consumer acceptance of electronic commerce: Integrating trust and risk with the technology acceptance model. *International Journal of Electronic Commerce, 7,* 3 (2003), 101–134.

98. Pennington, R.; Wilcox, H.D.; and Grover, V. The role of system trust in business-to-consumer transactions. *Journal of Management Information Systems, 19,* 1 (Summer 2002), 211–242.

99. Pereira, R.E. Influence of query-based decision aids on consumer decision making in electronic commerce. *Information Resource Management Journal, 14,* 1 (2001), 31–48.

100. Petty, R.E., and Cacioppo, J.T. *Communication and Persuasion: Central and Peripheral Routes to Attitude Change.* New York: Springer-Verlag, 1986.

101. Pittam, J. *Voice in Social Interaction: An Interdisciplinary Approach.* Thousand Oaks, CA: Sage, 1994.

102. Price, L.L., and Feick, L.F. The role of interpersonal sources in external search: An informational perspective. In T. Kinnear (ed.), *Advances in Consumer Research,* vol. 11. Provo, UT: Association for Consumer Research, 1984, pp. 250–255.

103. Qiu, L., and Benbasat, I. Do looks matter? A study of demographic embodiments of product recommendation agents in electronic commerce. Working Paper, Peking University, Beijing, 2008.

104. Reeves, B., and Nass, C. *The Media Equation: How People Treat Computers, Television, and New Media Like Real People and Places.* New York: Cambridge University Press, 1996.

105. Reynolds, K.E., and Beatty, S.E. Customer benefits and company consequences of customer–salesperson relationships in retailing. *Journal of Retailing, 75,* 1 (1999), 11–32.

106. Riegelsberger, J.; Sasse, M.A.; and McCarthy, J.D. Shiny happy people building trust? Photos on e-commerce Websites and consumer trust. In *Proceedings of the 2003 SIGCHI Conference on Human Factors in Computing Systems.* New York: ACM Press, 2003, pp. 121–128.

107. Riker, W.H. The nature of trust. In J.T. Tedeschi (ed.), *Perspectives on Social Power.* Chicago: Aldine, 1971, pp. 63–81.

108. Russo, J.E. Aiding purchase decisions on the Internet. Paper presented at the Winter 2002 SSGRR (Scuola Superiore Giuseppe Reiss Romoli) International Conference on Advances in Infrastructure for Electronic Business, Education, Science, and Medicine on the Internet, L'Aquila, Italy, July 29–August 4, 2002.

109. Sallnas, E.-L. Effects of communication mode on social presence, virtual presence, and performance in collaborative virtual environments. *Presence: Teleoperators and Virtual Environments, 14,* 4 (2005), 434–449.

110. Scherer, K.R. Vocal affect expression: A review and a model for future research. *Psychology Bulletin, 99,* 2 (1986), 143–165.

111. Scherer, K.R. Vocal communication of emotion: A review of research paradigms. *Speech Communication, 40,* 1–2 (2003), 227–256.

112. Short, J.; Williams, E.; and Christie, B. *The Social Psychology of Telecommunication.* London: John Wiley & Sons, 1976.

113. Sproull, L.; Subramani, M.; Kiesler, S.; Walker, J.H.; and Waters, K. When the interface is a face. *Human–Computer Interaction, 11,* 1 (1996), 97–124.

114. Straub, D.W. The effect of culture on IT diffusion: E-mail and fax in Japan and the U.S. *Information Systems Research, 5,* 1 (1994), 23–47.

115. Takeuchi, A., and Naito, T. Situated facial displays: Towards social interaction. In I. Katz, R. Mack, I. Marks, M.B. Rosson, and J. Nielson (eds.), *Proceedings of the 1995 SIGCHI Conference on Human Factors in Computing Systems.* New York: ACM Press, 1995, pp. 450–455.

116. Takeuchi, Y.; Takahashi, T.; and Katagiri, Y. Life-like agent design based on social interaction. Paper presented at the Sixth Pacific Rim International Conference on Artificial Intelligence, Melbourne, Australia, August 28–September 1, 2000.

117. Teo, T.S.H.; Lim, V.K.G.; and Lai, R.Y.C. Intrinsic and extrinsic motivation in Internet usage. *Omega: International Journal of Management Science, 27,* 1 (1999), 127–145.

118. Turkle, S. *Life on the Screen: Identity in the Age of Internet.* New York: Simon & Schuster, 1995.

119. Van der Heijden, H. Factors influencing the usage of Websites: The case of a generic portal in the Netherlands. *Information & Management, 40,* 6 (2003), 541–549.

120. Van der Heijden, H. User acceptance of hedonic information systems. *MIS Quarterly, 28,* 4 (2004), 695–704.

121. van Mulken, S.; Andre, E.; and Muller, J. The persona effect: How substantial is it? In B. Adelson, S. Dumais, and J. Olson (eds.), *Proceedings of HCI on People and Computers XIII.* Berlin: Springer, 1998, pp. 53–66.

122. Walker, J.H.; Sproull, L.; and Subramani, R. Using a human face in an interface. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.* New York: ACM Press, 1994, pp. 85–91.

123. Walther, J.B.; Slovacek, C.L.; and Tidwell, L.C. Is a picture worth a thousand words? Photographic images in long-term and short-term computer-mediated communication. *Communication Research, 28,* 1 (2001), 105–134.

124. Wang, W., and Benbasat, I. Trust in and adoption of online recommendation agents. *Journal of the Association for Information Systems, 6,* 3 (2005), 72–101.

125. Wang, W., and Benbasat, I. Recommendation agents for electronic commerce: Effects of explanation facilities on trusting beliefs. *Journal of Management Information Systems, 23,* 4 (Spring 2007), 217–246.

126. Wang, W., and Benbasat, I. Attributions of trust in decision support technologies: A study of recommendation agents for e-commerce. *Journal of Management Information Systems,* 24, 4 (Spring 2008), 249–273.

127. Xiao, B., and Benbasat, I. Customer decision support systems for e-commerce: Design and adoption of e-commerce product recommendation agents. *MIS Quarterly, 31,* 1 (2007), 217–309.

128. Yi, M.Y., and Hwang, Y. Predicting the use of Web-based information systems: Self-efficacy, enjoyment, learning goal orientation, and the technology acceptance model. *International Journal of Human–Computer Studies, 59,* 4 (2003), 431–449.

Appendix A: A Summary of Selected Empirical Evaluation Studies on Anthropomorphic Interface Agents

Paper	Independent variables and moderators (interface components)	Dependent variables and results	Application context
Walker et al. [122]	Synthesize talking face (face + text-to-speech voices) versus text Stern face versus neutral face	• Time: face > text; stern face > neutral face; Mistake: face < text; stern face < neutral face • Comments: face > text; stern face > neutral face • Likability: stern face > neutral face	Interview/questionnaire
Takeuchi and Naito [115]	Synthesized face versus arrow	• Engagingness: face > arrow • Usefulness: face < arrow	Poker game
Sproull et al. [113]	Synthesized talking face (synthesized face + text-to-speech voices) versus text	• Perceived intelligence: no difference • Social evaluation: talking face < text • Sociability: talking face < text • Arousal: talking face > text	Counselor interview
Koda and Maes [57]	Face (caricature or realistic) versus no face	• Intelligence: no difference • Likability: face > no face • Engagingness: face > no face • Comfortability: face > no face • Attention attraction: face > no face	Poker game
Lester et al. [70]	Animated pedagogical agents with various levels of communicative behaviors	• Believability: expressive agent > muted agent • Usefulness: expressive agent > muted agent	Tutoring system
van Mulken et al. [121]	No character versus with character (persona); all with text-to-speech voices Moderator: technical versus nontechnical information	• Engagingness: persona > no persona (technical info only) • Usefulness: persona > no persona (technical info only)	Tutoring system

(continues)

Paper	Independent variables and moderators (interface components)	Dependent variables and results	Application context
Nass et al. [91]	Ethnicity of an interface agent (simulated with real-human video): Caucasian versus Korean; subjects are all Korean males	• Value congruence: in-group (Korean agent) > out-group (Caucasian agent) • Social attractiveness: in-group > out-group • Trustworthiness: in-group > out-group • Argument quality: in-group > out-group	Collaborative choice-dilemma task
Bickmore and Cassell [16]	Task-oriented dialogue versus social dialogue generated by an embodied character with text-to-speech voice Moderator: introvert versus extravert subjects	• Trust: social dialogue > task dialogue (for extravert subjects only) • Engagingness: social dialogue > task dialogue	Information presentation
Link et al. [73]	Various pitch and speed of text-to-speech voice, various facial expression of an agent	• Feedback rating: no difference on voices; mouth curve up > mouth curve down	Learning feedback
McBreen and Jack [77]	Video versus a 3D talking head versus an image with quasidynamic facial expressions versus a still image versus a disembodied voice; all with human voice Moderator: human-like versus cartoon-like	• Overall rating (human-like agents): video > voice only > still image > image with facial expressions > 3D talking head • Overall rating (cartoon-like agents): no difference	Online shopping
Moreno et al. [82]	With animated character versus without character Text versus voice	• Knowledge transfer: character > no character • Study interests: character > no character; voice > text • Retention: voice > text	Tutoring system
Moundridou and Virvou [83]	With embodied (face + text-to-speech voice) agent versus without embodied agent (no face + text)	• Enjoyment: with face > without face • Ease of use: with face> without face • Usefulness: with face > without face	Tutoring system

Baylor and Ryu [12]	With image versus without image Static image versus animated image All with both voice and on-screen text	• Credibility: with image > no image • Person-like: animation > static • Instructor-like: animation > static • Engagingness: animation > static	Tutoring system
Nowak and Biocca [94]	No image versus less-anthropomorphic image versus highly anthropomorphic image	• Telepresence: low-anthropomorphic image > high-anthropomorphic image • Copresence: low-anthropomorphic image > high-anthropomorphic image • Social presence: low-anthropomorphic image > high-anthropomorphic image	Virtual environment
Keeling et al. [55]	Realistic human versus cartoon human versus carton animal	• The implementation of an embodied agent would be preferred when it (1) resembled a person found in the corresponding offline situation, (2) was associated with the site as a whole, (3) appealed to likely users of the site, and (4) represented a product sold on the Web site	Online shopping
Baylor [11]	Gender and ethnicity of the agent Cartoon-like agent versus realistic agent Moderator: agent's role as expert versus motivator	• Subjects tend to choose an agent of the same ethnicity when given a chance to choose • Usefulness/interesting/satisfaction: male agents > female agents • Intelligence, knowledge, expert-like: male > female (motivator agent) • Competence: male > female (expert agent) • Enjoyment/enthusiasm/motivation: African American > Caucasian (motivator agent)	Tutoring system

(continues)

Paper	Independent variables and moderators (interface components)	Dependent variables and results	Application context
Berry et al. [15]	Human versus animated agent (face + text-to-speech voice) versus text-to-speech voice-only versus text-only	• Ease of understanding: text-only > agent, TTS-only, human • Trustworthiness: TTS-only, text-only, agent > human • Helpfulness: human, agent > TTS-only • Likability: human > TTS-only	Information presentation/ persuasion
Cowell and Stanney [26]	Demographic embodiments of a humanoid agent Nonverbal interaction styles	• Users prefer to interact with characters that match their ethnicity and are young looking (when only the picture of an agent was presented) • Credibility: agent with trustworthy facial nonverbal behavior > agent without nonverbal behavior; no difference on postures/gestures	Collaborative task (photo sorting)
Holzwarth et al. [49]	Avatar versus no avatar Attractive avatar versus expert avatar	• Avatar sales agent leads to more satisfaction with the retailer, a more positive attitude toward the product, and a greater purchase intention • Attractive avatar is more effective at moderate levels of product involvement; expert avatar is more effective at high levels of product involvement	Online shopping

Appendix B: Measurement Items

Construct (source)	Items
Social presence [38]	I felt a sense of human contact in the agent.
	I felt a sense of personalness in the agent.
	I felt a sense of human warmth in the agent.
	I felt a sense of sociability in the agent.
	I felt a sense of human sensitivity in the agent.
Trusting beliefs [125]	The agent was competent in recommending digital cameras.
	The agent performed its role of recommending digital cameras very effectively.
	Overall, the agent was capable of providing suitable digital camera recommendations.
	In general, the agent was very knowledgeable about digital cameras.
	I believe that the agent's dealings with me were in my best interest.
	The agent's dealings with me felt like that it would do its best to help me.
	The agent's dealings with me felt like that it was interested in my well-being, not someone else's.
	I believe the agent's recommendations to me were truthful.
	I would characterize the agent's dealings with me as honest.
	The agent appeared to be unbiased.
Perceived enjoyment [59]	Interacting with the recommendation agent is enjoyable.
	Interacting with the recommendation agent is exciting.
	Interacting with the recommendation agent is pleasant.
	Interacting with the recommendation agent is interesting.
	Interacting with the recommendation agent is fun.
Perceived usefulness [28]	PU1: Using the recommendation agent can improve my shopping performance.
	PU2: Using the recommendation agent can increase my shopping productivity.
	PU3: Using the recommendation agent can increase my shopping effectiveness.
	PU4: I found using the recommendation agent useful.
Agent usage intentions [28]	AUI1: If I have access to the system, I intend to use the agent for my next digital camera purchase.
	AUI2: If I have access to the system, I predict I would use the agent for my next digital camera purchase.
	AUI3: If I have access to the system, I plan to use the agent for my next digital camera purchase.

Appendix C: Results of the Mediation Tests

Our research model was also examined with the mediation tests suggested by Baron and Kenny [8]. Consistent with H2, H3, H5, and H7, trust and perceived enjoyment fully mediate the effect of social presence on perceived usefulness. The regression analyses in Table C1 show that social presence was a significant determinant of trust ($p < 0.001$) and perceived enjoyment ($p < 0.001$), and the significant effects of social presence on perceived usefulness became nonsignificant when trust and perceived enjoyment were controlled for.

The regression analyses in Table C1 also reveal that trust and perceived enjoyment were significant determinants of perceived usefulness ($p < 0.001$ for both), and the significant effects of trust and enjoyment on usage intention became nonsignificant (for enjoyment) and less significant (for trust) when perceived usefulness was controlled for. This implies that perceived usefulness partially mediates the effect of trust and fully mediates the effect of perceived enjoyment on usage intention.

Table C1. Regression Results

Dependent variable	R^2	Independent variable	β	SE_β	Significance level
Hypotheses 2, 3, 5, 7: Trust and perceived enjoyment mediate the effect of social presence on perceived usefulness					
Trust	0.147	Social presence	0.384	0.042	< 0.000
Perceived enjoyment	0.396	Social presence	0.630	0.052	< 0.000
Perceived usefulness	0.149	Social presence	0.386	0.060	< 0.000
Perceived usefulness	0.427	Social presence	–0.019	0.066	0.804
		Trust	0.413	0.092	< 0.000
		Perceived enjoyment	0.392	0.075	< 0.000
Hypotheses 4, 5, 6, 7: Perceived usefulness partially mediates the effect of trust and fully mediates the effect of perceived enjoyment on usage intentions					
Usage intention	0.343	Trust	0.882	0.142	< 0.000
		Perceived enjoyment	0.399	0.097	< 0.000
Perceived usefulness	0.427	Trust	0.582	0.090	< 0.000
		Perceived enjoyment	0.370	0.062	< 0.000
Usage intentions	0.555	Trust	0.363	0.131	0.006
		Perceived enjoyment	0.069	0.088	0.439
		Perceived usefulness	0.893	0.101	< 0.000