Telematics and Informatics 48 (2020) 101344

Contents lists available at ScienceDirect

Telematics and Informatics

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

Language, modality, and mobile media use experiences: Social responses to smartphone cues in a task-oriented context

Kun Xu

Department of Telecommunication, College of Journalism and Communications, University of Florida, Gainesville, FL 32611, United States

ARTICLE INFO

Keywords:
The computers are social actors paradigm
Social cues
Smartphone
Mobile media use experiences
Language cues
Modality cues

ABSTRACT

With people's growing use of virtual agents and voice assistants on smartphones, researchers have pointed out that mobile phones are not only acting as intermediaries that connect users from different places, but also communication terminals that present different combinations of social cues. This study applies the Computers are Social Actors paradigm in human-phone interaction and postulates that compared to non-anthropomorphic language and text cues, anthropomorphic language and vocal cues will have more effects on users' social responses to smartphones. This study also explores the role of individual differences in users' social responses to smartphones. Based on a lab experiment using a between-subjects factorial design, the study suggests that although anthropomorphic language and voice-based information did not have main effects on users' social responses, people's mobile media usage and intensive phone use interacted with these cues in their social responses to the smartphones. In addition, this study implies that it is the combination of social cues, individual differences, and communication contexts that contributes to people's social reactions to the smartphones. The findings of the study can inform user interface design and precipitate further discussion about the ethical issues in human-phone interaction.

1. Introduction

As a tool for "anytime, anywhere connectivity" (Ito, 2005, p. 131), mobile phones have diffused at an unprecedented speed in the past decades. The number of mobile phone users reached 50% of the US residents within only 14 years (Wei, 2013). Mobile broadband also increased by 45% annually from 2010 to 2014 (Mihailidis, 2014). Since 2015, the number of mobile phone users has increased from 4.15 billion to 4.78 billion worldwide (Statista, 2019). In 2018, 95% of the US adults owned at least one mobile phone (Pew Research Center, 2018), and mobile phones afforded more than half of the global website traffic (Deyan, 2019).

Along with the growing number of users, mobile phones have evolved from merely portable devices that connect people from different places to smart devices that feature multi-structured platforms and multi-faceted modalities (Zhong, 2013). Today mobile phone users can send emails, browse the Internet, play video games, listen to podcasts, read books, and use thousands of mobile applications available in the market (Scolari et al., 2012). The rapid growth and expansion of mobile networks and smartphones demonstrates that phones are advancing in both mobility and computing power (Humphreys, 2013). Given the increasing usage of smartphones and mobile apps, investigating why mobile users are attracted to these technologies and how they evaluate the information from them can help us better understand the role of mobile technologies in our daily lives.

Whereas past works on mobile media have primarily focused on how users treat these devices as communication channels (Campbell, 2013), more recent studies have begun to concentrate on the direct interactions between users and phones. For instance,

E-mail address: kun.xu@ufl.edu.

https://doi.org/10.1016/j.tele.2020.101344
Received 23 June 2019; Received in revised form 18 December 2019; Accepted 11 January 2020
Available online 16 January 2020

Carolus et al. (2019a) found that mobile phones not only serve technical purposes but also function as digital companions of users. Carolus and colleagues (2019b) further found that smartphones that used polite language were perceived to be friendlier than those that used impolite language. As the diffusion of the smartphones including mobile virtual assistants have allowed users to more frequently react to the vocal, haptic, and language cues of the phones, understanding how users perceive, evaluate, and interact with them would help researchers anatomize the relationships between humans and technologies in the current digital age. Thus, this study applies the Computers are Social Actors (CASA) paradigm and examines the effects of two social dimensions of smartphones: language and modality. Additionally, considering that CASA researchers have called for studies on the role of individual differences in users' social responses to technologies (Nass and Moon, 2000), this study probes how individuals' mobile media use experiences add to the effects of the social dimensions on individuals' psychological responses.

The findings of the current study can make the following theoretical contributions. First, the findings would add more evidence to the expansion of the CASA paradigm to the field of human-phone interaction. Second, the study would parse out the interaction between individual differences and technology features in evoking users' social responses to smartphones. Third, the comparisons between two pairs of social cues are expected to render a hierarchy of social dimensions that have different power in activating users' reactions. These theoretical advancements could further be translated into applied knowledge in areas such as mobile interface design, education, and mobile intervention.

2. Literature review

In this section, a brief review about mobile media studies and the effects of social cues is introduced. Then language cues and modality cues are introduced as the two major factors in the study. Toward the end we discuss users' social responses and individual differences, followed by our research questions.

2.1. Mobile media and the CASA paradigm

Prior literature about people's use of mobile media has covered topics including text messaging, uses and effects of mobile phones, and identity formation (Katz, 2007). Wei (2008) applied the uses and gratifications approach and discovered that people used mobile phones for passing time, social interaction, reassurance, instrumentality, and communication facilitation. Some scholars approached smartphone research through the lens of geographic information (de Souza e Silva, 2013; Frith, 2014; Jethani and Leorke, 2013). It has been found that locative mobile media have led to users' adoption of geographic information systems and mapping tools (Dredze et al., 2013; Jethani and Leorke, 2013; Sutko and de Souza e Silva, 2010). Researchers have further looked at the social and cultural aspects of mobile phones. Nardi et al. (2000) found that instant messaging via mobile phones not only facilitated information exchange, but also shaped an outeraction process where users negotiate availability and maintain social connections. More recently researchers have examined the cultural discourses about mobile phones and suggested that integration and disintegration are the two major narratives that affect users' perception of the mobile devices, designers' vision of the technology innovation, and scholars' framing of their questions (Harmon and Mazmanian, 2013).

Despite prior research on uses and effects of mobile media, location-based services, and sociocultural aspects of mobile phones, limited research has explained how smartphone as a device can be perceived as a social entity (Carolus et al., 2019b; Kim, 2014). Although two studies discovered that the presence of a mobile phone between two face-to-face communication partners reduced individuals' engagement in conversations and lowered partners' intimacy and mutual understanding (Misra et al., 2016; Przybylski and Weinstein, 2013), still more research is needed to understand how phones are perceived as social agents rather than mere machines.

To understand users' social responses to smartphones, the CASA paradigm (Nass et al., 1994; Reeves and Nass, 1996) is applied as a theoretical framework in this study. It suggests that users transfer the social scripts of human–human communication to human–computer interaction. For instance, computer users apply etiquette norms to computers, use gender stereotypes to evaluate computers, and attribute socialness to computers rather than the programmers of the computers (Nass et al., 1994). Based on a series of experimental studies on computers and televisions, Reeves and Nass (1996) used the media equation to describe how individuals treat media as real people. More recently, the media equation has been applied to users' interactions with virtual agents and social robots. As some examples, Kim, Han, Jung, and Lee (2013) suggested that users reported greater enjoyment and involvement when they saw the robot NAO's motion of arms and head. Lee, Peng, Jin, and Yan (2006) manipulated the task quantity and difficulty in users' interaction with a social robot Sony AIBO. They found that users were more attracted to a growing robot than to a mature one. In the context of human-agent interaction, Jung et al. (2014) found that an embedded tutor agent on a programming interface was perceived as more humanlike and attractive than an external agent embedded in a separate device. Liew and Tan (2018) found that specialist virtual agents improved the credibility of an online shopping website compared to generalist virtual agents.

To recognize the features of technologies that evoke users' social responses to computer technologies, Nass (2004) proposed a list of social cues that trigger users' etiquette responses to computers. These cues include language use, human and synthetic voice, face, emotion, interactivity, and social roles. Fogg (2002) also identified five categories of cues that activate users' psychological processing: physical cues (e.g., face, eye), psychological cues (e.g., empathy, humor), language cues (e.g., language recognition), social dynamics (e.g., praise, reciprocity), and social roles (e.g., teammates, teachers).

Whereas researchers have recognized the significance of cues and found that greater number of cues would trigger stronger social responses (e.g., Bartneck et al., 2008; Breazeal, 2003; Duffy and Zawieska, 2012), the comparisons within a single category of social cues have been largely ignored. That is, little research has examined the quality of cues and investigated the different power of

elements under one broad category of cues. For example, voice can be classified into volume, speed, and pitch (Chang et al., 2018; Joosse et al., 2013). Language can be sorted into familiar vs. horrific styles (Kim et al., 2013), or agreeable vs. disagreeable ones (Martínez-Miranda et al., 2018). These subtle elements can all have different impact on users' perception and attitudes toward technologies. Thus, a more detailed analysis of different properties of social cues should be conducted to understand how the quality of each single cue contributes to users' psychological processing of media (Lombard, 2018).

This study specifically centers on two sets of cues: language cues and modality cues. Making changes to the language of smartphones has been one of the approaches to enhancing user experience. Both Google's Material Design guidelines (2017) and Apple's iOS Human Interface guidelines (2017) have acknowledged the importance of human-centered language in designing voice-based information. Similarly, modality cues have been found to affect the perceived quality and credibility of user interfaces (Sundar et al., 2015). Oh et al. (2013) compared page flipping and double clicking on a website and found that digital page flipping increased the usability and engagement of the website.

Along with the necessity to inquire into the modality and language cues, it remains unknown whether responses to these cues are contingent upon any individual differences or contexts (Nass and Moon, 2000). In an attempt to update the CASA paradigm, Xu and Lombard (2016) proposed that it is not only the social cues, but also the combination of social cues, individual factors, and communication contexts that lead people to perceive media as social actors. Although initially Nass and Moon (2000) found no effects of expertise or gender on users' social responses to computers, succeeding studies noted that those who tended to be more analytical and critical were less likely to accept computers' suggestions when they were flattered by the computers (Lee, 2010). Research also revealed that age and personalities could affect users' perception and acceptance of robots' personalities (Woods et al., 2007). Therefore, given the lack of research on the quality of social cues, the limited application of the CASA paradigm to human-phone interaction, and the inconsistencies of the influence of personal factors on users' social responses, this study compares two pairs of social cues and assesses whether the paradigm can be expanded to users' interactions with phones and how individual differences such as mobile media use experiences in a task-oriented context affects the expansion of the theory.

2.2. Language and modality cues

Language is defined as "the expression of human communication through which knowledge, belief, and behavior can be experienced, explained, and shared" (NIH, 2017). Though language can be reflected in many forms, this study centers on written language and spoken language to match the examination of the modality cues.

Prior research has suggested that the language impact rests on language cues. Adkins and Brashers (1995) compared powerful language versus powerless language (i.e., use of hedges, polite forms, hesitations in messages) in the context of computer-mediated communication. They found that powerful language increased the perceived trustworthiness and attractiveness of the information senders. In the context of human–computer interaction, Nass et al. (1995b) focused on the differences between dominant language (i.e., assertive, strong, confident language) and compliant language (i.e., hesitant, conservative, and indefinite language). They found that the participants who identified themselves as dominant tended to perceive the computer with the dominant language as more intelligent, more attractive, and more convincing. Fogg (2002) also found that the participants who received complimentary language from a computer felt better about themselves, performed better in a task, and were more willing to work with the computer.

Another way to classify language is to review the anthropomorphic cues designed into language (Araujo, 2018). Anthropomorphic cues are the elements that have human characteristics. Past work has demonstrated the different effects of anthropomorphic images. For example, Gong (2008) compared facial images with different levels of human-likeness. He found that the more anthropomorphic the images were, the more likely the participants perceived the characters in the images as positive, trustworthy, homophily, and persuasive.

While previous literature has shown that language styles would impose their influence on users' social responses to computers (Nass et al., 1995b), limited research has directly compared the differences between anthropomorphic language and non-anthropomorphic language in users' social reactions to technologies, especially smartphones. In prior research, active voice and personal nouns were identified as anthropomorphic linguistic cues (Sah and Peng, 2015). Sah and Peng (2015) found that using these cues on a website evoked users' greater social perception. Vinciarelli et al. (2009) listed voice quality (e.g., pitch, tempo, frequency), linguistic vocalizations (e.g., "ehm", "aha", "hmm"), and non-linguistic vocalizations (e.g., laughter, whispers) as cues that can contribute to the social perception of a message. Similarly, Zhao (2003) considered deliberate misspelling, exaggeration of tones, and emoticons as important factors in people's sense of co-presence. Self-referential statements presented by a computer can be another influential factor in users' social responses (Nass and Brave, 2005). It was found that participants reported first-person voice to have higher levels of humanness than passive voice.

Overall, these studies have indicated that the more anthropomorphic linguistic cues that are designed into messages, the stronger users' social responses will be. Based on the relationship, the current study extrapolates that in human-phone interaction, compared to non-anthropomorphic language, anthropomorphic language will evoke users' stronger social responses to smartphones.

In addition to language cues, voice and text are perceived as the major modalities of mobile phones. Media users are sensitive to human voice as a social cue of technologies (Nass and Steuer, 1993). Human voice can carry a speaker's emotions and attitudes and thus is perceived as more natural and powerful in human–computer interaction (Nass and Brave, 2005). As an example, Takayama et al. (2009) investigated how voice location would affect users' responses to a social robot. They found that when a robot's voice was sent out from a separate control box instead of the robot's own physical entity, users were more likely to accept criticism. Takayama et al. (2009) explained that the dissociation between the source of voice and the physical embodiment of the robot attenuated the perceived aggressiveness of the robot and users automatically and intuitively reacted to the voice rather than the robot per se. The

study showed that users were sensitive to voices especially when the content of the voices was associated with negative comments.

Some scholars have already compared the different effects of voice and text. Crooks et al. (2012) found that when participants were not experiencing high cognitive load, text-based information was more likely to enhance users' free recall performance and message comprehension than speech-based information. But there are also studies advocating the primacy of voice over text in evoking people's social perception. Burgoon and colleagues (2000) argued that compared to voice, text is perceived as neutral and less likely to evoke users' attributions of human characteristics. Sproull et al. (1996) corroborated this perspective and reported that when participants interacted with a voice-based user interface rather than a text-based interface, they felt more aroused and more positive toward the user interface. Considering that human voice has special acoustic properties that are familiar to humans (Nass and Steuer, 1993), this study postulates that compared to text, voice will arouse users' stronger social responses to smartphones.

2.3. Social responses

As indicators of users' social responses to media technologies, previous research has investigated the role of users' social presence, social attitudes, and social behavior in human-technology interaction. While social presence has been conceptualized from various perspectives ranging from "the subjective quality of the medium" (Short et al., 1976, p. 66) to "sense of being with others" (Zhao, 2003, p. 450), this study adopts Lombard and Ditton's (1997) conceptualization, where social presence could occur when users respond to the social cues presented by the media themselves. As this study examines users' direct and social responses to smartphones, this conceptualization best captures the current research context.

The effects of social cues on users' social attitudes have been associated with how users feel attracted to the media technologies and how users trust the information provided by these media technologies (Jung and Lee, 2004; Jung et al., 2014; Nass et al., 1996). In addition, conformity can manifest users' social behavior in their interactions with media technologies. Conformity refers to one's actual behavioral change to reach consensus between communicators (Lee and Nass, 2002). Prior research has shown that even when interacting with computers or robots, users may conform to the suggestions from these technologies (Nass et al., 1996; Takayama et al., 2009). According to the CASA paradigm, if users treat phones as social actors, users should demonstrate similar behavioral change as they do in interpersonal communication. Therefore, based on prior literature on the CASA paradigm, social perception, social attitudes, and social behavior, the following hypotheses are proposed.

***H1*:** Compared to one with non-anthropomorphic language, a mobile phone with anthropomorphic language will lead to greater levels of social responses, which include a) social presence, b) perceived attraction of the phone, c) perceived trustworthiness of the phone, d) users' conformity intention, and e) users' conformity to the phone.

H2: Compared to one with text-based information, a mobile phone with voice-based information will lead to greater levels of social responses, which include a) social presence, b) perceived attraction of the phone, c) perceived trustworthiness of the phone, d) users' conformity intention, and e) users' conformity to the phone.

2.4. Interindividual differences

Different extent of mobile media use may lead to different types and strength of social responses. Although early HCI studies showed that technology use experiences did not affect users' social responses to desktop computers (Nass et al., 1995a), studies thereafter demonstrated the opposite evidence that experienced computer users reported more confidence and more trust in a computer when they were complimented by the computer (Johnson et al., 2004). Xu (2019) also found that prior robot interaction experiences moderated the relationship between a social robot's vocal cues and users' social attitudes. Hence, individuals' technology use experiences may play a role in users' social responses to technologies.

In this study mobile media use experiences are inquired from two perspectives: how much time people spend on mobile media and how intensive they use mobile media. The former has been conceptualized as mobile media usage in prior research and was found to affect individuals' emotional and physical status (Liu, 2010; Panda and Jain, 2018). The latter has been referred to as intensive mobile phone use or power use (Davis, 2002; Derks and Bakker, 2014; Zhong, 2013). Intensive phone users often exploit the mobile devices to the fullest extent (Schlosser, 2002). They use the mobile devices more creatively and efficiently than ordinary users (Zhong, 2013). Compared to others, their mobile media usage may not be high, but they utilize more default features of the devices (Zhong, 2013). Derks and Bakker (2014) found that for intensive phone users, engaging in recovery activities could lead to mental detachment from work. But when facing greater life interference, intensive phone users felt more exhausted than non-intensive phone users. As these two types of mobile media use experiences could both influence users' attitudinal and behavioral reactions, we propose the following research questions.

RQ1: How will users' mobile media usage interact with the language cues and the modality cues of the phones in predicting users' social responses?

RQ2: How will users' intensive mobile phone use interact with the language cues and the modality cues of the phones in predicting users' social responses?

Furthermore, researcher have not reached consensus on whether users' gender would affect their social responses to technologies. Although Nass and Moon (2000) suggested that social responses to computers were not confined to gender differences, females were found to be more sensitive to computers' flattery than males (Lee, 2008). They were also less likely to accept computers taking personal roles such as babysitters, judges, and psychiatrists than males (Nass et al., 1995a). Xu (2019) also found that males had more intention to use a social robot than females if the robot could demonstrate humanlike gestures. In the context of human-phone interaction, gender differences in people's mobile use habits were not clear. While Okazaki and Hirose (2009) found that females

obtained more satisfaction from mobile Internet use than males, DeBaillon and Rockwell (2005) pointed out that the gaps between males and females in mobile phone use had substantially narrowed. Therefore, this study continues to investigate the gender differences in human-phone interaction.

RQ3: How will users' gender interact with the language cues and the modality cues of the phones in predicting users' social responses?

3. Method

3.1. Participants

A total of 111 participants were recruited from a public university in Northeast U.S. to voluntarily participate in a lab experiment. One participant's responses were excluded due to unexpected interruption during the experiment. Participants were rewarded with extra credit for the participation. Half of them are males and the other half are females (50%). The average age was 20.44 years old ($M = 20.44$, $SD = 3.41$). On average, participants spent 6.25 h per day on their mobile phones ($M = 6.25$, $SD = 4.32$).

3.2. Research design

A voice assistant mobile app was developed and programmed for the experiment. The mobile application was installed on a 4.87-inch tall, 2.31-inch wide, and 0.30-inch deep "Space Gray" iPhone 5S. The phone weighs 3.95 oz and has a 4-inch diagonal Retina display with multi-touch functions. On the user interface of the application, the background color was set as white and any given text information was shown in black. Three blue buttons were designed at the bottom of the interface which were labeled as "PREV", "PLAY" (or "DISPLAY"), and "NEXT."

The experiment adopted a 2 × 2 between-subjects factorial design. Participants were randomly assigned to one of the four conditions: voice with anthropomorphic language ($n = 28$), voice with non-anthropomorphic language ($n = 27$), text with anthropomorphic language ($n = 27$), and text with non-anthropomorphic language ($n = 28$).

To design anthropomorphic language, anthropomorphic elements identified in prior research were included in the manipulation of the messages. The messages used active tense (Sah and Peng, 2015), self-referential statements (e.g., "I," "my") (Nass and Steuer, 1993), informal conversation speech (e.g., "it's like…," "well…") (Vinciarelli et al., 2009), and exaggeration of tones (e.g., "I like it soooo much," "Super!!!") (Zhao, 2003). Comparatively, non-anthropomorphic language used more passive tense, non-self-referential statements, and formal statements. To manipulate the voice conditions, a pre-recorded female human voice was designed into the user interface. Participants interacted with the phone and listened to the messages. In the text conditions, the human-voiced messages were converted to texts and participants read the messages on the phone.

3.3. Procedures

IRB approval was obtained before the experiment. After participants arrived at the lab and read the consent form, they first filled out some demographic information via the survey software Qualtrics on a laptop. Then they were told to work on a hypothetical survival scenario "Lost at Sea" (Humber, 2017). Survival problems have been used in prior research on users' social responses to computers and social robots (Nass et al., 1996; Takayama et al., 2009). Participants were provided with a scenario in which they had rented a yacht with three friends. A fire broke out on the yacht and they lost the crew. The yacht was damaged and gradually sinking. Participants were told that the location of the yacht was unclear. They only knew that they might be hundreds of miles away from the nearest landfall. Participants were told that they had already secured a rubber life raft and a box of matches. In addition, there were 15 reserved items undamaged. Due to the limited space in the rubber life raft, they can only take eight of the 15 items with them on the life raft. They had already reached consensus with their friends on taking a 25-liter container of portable water with them, but they need to make decisions on the priorities of the other 14 items.

Seven pairs of items were presented to the participants on the laptop. They included a shaving mirror vs. a 10-liter can of oil gasoline mixture, a case of army rations vs. two boxes of chocolate bars, and other five pairs of tools. The participants were asked to make decisions on which tool to retain from each pair of items.

After participants made their initial choices, they were notified that they would interact with a mobile phone. The researcher turned on the mobile app and gave the mobile phone to the participants. This is to minimize the possibility that participants felt they were interacting with the app rather than the phone. Then participants were instructed to press "PLAY" or "DISPLAY" to receive further information.

In the voice with anthropomorphic language condition, they heard a voice saying "Hello, I'm the XX Lab (*anonymity*) application. I'm sooo happy to meet you. I'm here to help you with the Lost at Sea scenario! I found seven hints that I can share with you! On each page you can press NEXT to receive the next hint. And you can press PREV if you would like to receive the previous hint. You can press PLAY to let me present the hint. Now, let's go to the first hint!" When the participants pressed "NEXT," the application provided its first suggestion on the Lost at Sea problem. The hint was "I have the first hint for you! You may wonder whether a shaving mirror or a 10-liter can of oil/gasoline mixture would help more if you are lost at sea. Well, I think a shaving mirror is more important. You know why? Because in sunlight a mirror can generate five to seven million candlepower of light. The reflected sunbeam can be seen beyond the horizon!" The participants can press "NEXT" to listen to the next hint. They can also press "PLAY" if they want to hear the message again.

The participants assigned to the text with non-anthropomorphic language condition were asked to read the messages on the user interface of the mobile application. For example, after they pressed the "DISPLAY" button, they saw the following text-based messages. "This is the M.I.N.D. Lab application. This application contains some information about the Lost at Sea scenario. Seven hints are installed in the application. Press NEXT to receive the next hint. Press PREV to receive the previous hint. Press DISPLAY to let it show the hint. Now go to the first hint." After the participants pressed "NEXT," they would receive the first suggestion. The message was "This is the first hint. Between a shaving mirror and a 10-liter can of oil/gasoline mixture, the mirror is more important. In sunlight, a mirror can generate five to seven million candlepower of light. The reflected sunbeam can be seen beyond the horizon."

The voice with non-anthropomorphic language condition used the same messages as the text with non-anthropomorphic language condition. The text with anthropomorphic language condition used the same messages as the voice with anthropomorphic language condition.

Participants were asked to finish reading or listening to all the eight messages from the phone and then make their final decisions on the priorities of the survival items. They were told that their answers would be judged based on the differences between their final choices and a survival expert's recommendations. At last they were asked to complete the rest of the questionnaire items including both closed-ended measures and open-ended questions about whether and why they conformed to the suggestions from the phone. The whole process took about 30 min. Participants were provided with debriefing forms when they left the lab.

3.4. Measures

The measure of social presence ($M = 6.39$, $SD = 1.89$, $\alpha = 0.80$) was adapted from previous measures of social presence in HCI contexts (Lee and Nass, 2005; Lee et al, 2006). Participants were asked to report their feelings on a Likert-type scale with seven 10-point items (1 = not at all, 10 = very much). The items include "How much did you feel as if the mobile phone was talking to you?" and "How much did you feel as if you and the mobile phone were communicating with each other?"

The measure of perceived social attraction ($M = 4.24$, $SD = 2.49$, $\alpha = 0.89$) was adapted from previous measures of social attraction (Lee et al., 2006; McCroskey and McCain, 1974). Participants were asked to report disagreement or agreement on a 10-point Likert-type scale with four items (1 = strongly disagree, 10 = strongly agree). The items include "I think I could establish a personal relationship with the mobile phone" and "I think I could have a good time with the mobile phone."

The measure of perceived trustworthiness ($M = 7.93$, $SD = 1.63$, $\alpha = 0.81$) was adapted from previous measures of trust (Gong, 2008; Gong and Nass, 2007). Participants were asked how they felt about the phone on four items. The four items use a 10-point semantic-differential scale. Examples of these items include "Untrustworthy - trustworthy" and "unreliable – reliable."

Conformity intention ($M = 8.31$, $SD = 1.60$, $\alpha = 0.86$) was adapted from the Kim and Park (2011) measure of conformity intention and Nass et al.'s (1996) measure of openness to influence. Participants were asked to report on seven 10-point Likert-type items (1 = strongly disagree, 10 = strongly agree). Examples of the items include "I was willing to follow the opinions from the mobile phone" and "I was dependent on the mobile phone's advice."

Actual conformity ($M = 9.54$, $SD = 1.46$) was adapted from the measures of behavioral change (Takayama et al., 2009; Lee, 2006). Conformity was operationalized as the change of participants' choices. It was measured by calculating the percentage of the number of changes that the participants made to match the suggestions from the phone and the number of the inconsistencies between the participants' initial choices and the suggestions from the mobile phone. The percentage was converted to a 10-point system.

Mobile media usage was adapted from Nass et al.'s (1995a,b) measure of technology use experiences. Participants were asked to indicate how many hours they spent using their mobile phone on a typical day and how many hours they spent using their electronic tablets on a typical day. Responses to the measure were summed and averaged to form the index ($M = 6.52$, $SD = 4.52$).

The measure of intensive phone use ($M = 5.98$, $SD = 0.96$, $\alpha = 0.69$) was adapted from previous measures (Derks and Bakker, 2014; Zhong, 2013). Participants were asked to report on a seven-point Likert-type scale with three items (1 = strongly disagree, 7 = strongly agree). Examples of the items include "I make good use of most of the features available in my mobile media device" and "I use my mobile media device intensively."

For manipulation check, participants were asked to indicate their perception of the language on the mobile phone. The measure was adapted from Sah and Peng's (2015) measure of anthropomorphic linguistic cues and used two 10-point semantic-differential items ($M = 6.50$, $SD = 2.72$, $\alpha = 0.87$). The items include "impersonal – personal" and "formal – conversational." A two-way ANOVA with the modality and the language as the independent variables and participants' perception of language as the dependent variable was conducted. Those in the anthropomorphic language conditions ($M = 7.59$, $SD = 1.95$) reported the language to be more personal and conversational than those in the non-anthropomorphic language conditions ($M = 5.47$, $SD = 3.00$), $F(1, 103) = 20.76$, $p < .001$, partial $\eta^2 = 0.17$. The manipulation of the language was successful.

3.5. Data analyses

For data management, univariate and multivariate outliers were examined using Box plots, Stem-and-leaf plots, extreme values, and Mahalanobis distance. As the whole dataset includes less than 5% of missing data for each variable, listwise deletion was used as they could be considered missing at random (Schafer, 1999). The variables of mobile media usage, intensive phone use, conformity intention, and actual conformity were log transformed as they were skewed in distribution. Three-way ANOVAs with the language cues, the modality cues, and participants' gender as independent variables were used to test H1, H2, and RQ3. To answer RQ1 and RQ2, the statistical tool Process macro was used to run the regression-based moderation analyses.

Table 1
The effects of language and modality cues on users' social responses (H1 & H2).

	Anthropomorphic language	Non-Anthropomorphic language	Main effects	Voice	Text	Main effects
	M (SD)	*M (SD)*	*F*	*M (SD)*	*M (SD)*	*F*
Social presence	6.56 (1.77)	6.22 (2.00)	0.86	6.46 (1.96)	6.32 (1.84)	0.20
Perceived attraction	4.64 (2.70)	3.85 (2.22)	2.65	3.94 (2.49)	4.54 (2.48)	1.54
Perceived trustworthiness	8.14 (1.65)	7.72 (1.61)	1.64	7.97 (1.91)	7.89 (1.33)	0.11
Conformity intention	8.24 (1.66)	8.37 (1.56)	0.16	8.26 (1.63)	8.36 (1.59)	0.05
Conformity	9.34 (1.83)	9.75 (0.93)	1.81	9.54 (1.26)	9.55 (1.65)	0.01

M: mean value. SD: standard deviation. df: degrees of freedom. df1 = 1, df2 = 99.

4. Results

H1 predicted that anthropomorphic language would be more likely than non-anthropomorphic language in evoking users' social responses to mobile phones. Results suggested that there were no significant differences between the language cues in evoking social presence, perceived attraction of the phone, perceived trustworthiness of the phone, conformity intention, and users' actual conformity. Therefore, *H1* was rejected. However, as it was found in the manipulation check that language cues had positive effects on participants' perception of language, post hoc analyses were conducted to test whether perception of language could further mediate the relationships between language cues and users' social responses. The mediation analyses can demonstrate how language cues as an antecedent variable links to users' social responses through users' perception of language. Using the Process macro, results suggested that there were indirect effects of the anthropomorphic language on participants' social presence through how personal and conversational they perceived the language to be, $B = 0.58$, 95% *CIs* [0.23, 1.02]. There were also indirect effects of the anthropomorphic language on perceived attraction of the phone, $B = 0.48$, 95% *CIs* [0.11, 0.93], and perceived trustworthiness of the phone, $B = 0.38$, 95% *CIs* [0.10, 0.72].

H2 predicted that voice would be more likely than text in evoking users' social responses to smartphones. Results suggested that there were no significant differences between voice- and text-based messages in predicting social presence, perceived attraction of the phone, perceived trustworthiness of the phone, users' conformity intention, and their actual conformity. Thus, *H2* was rejected. The results of H1 and H2 were shown in Table 1.

RQ1 asked how users' mobile media usage would interact with the language or the modality cues in predicting users' social responses. Moderation analyses revealed that although mobile media usage did not interact with language cues in predicting perceived trustworthiness of the phone, $B = -2.38$, $p = .09$, 95% *CIs* [−5.13, 0.38], the Johnson-Neyman technique suggested that when the value of log transformed mobile media usage was lower than 0.70, there was a significantly positive relationship between anthropomorphic language and the perceived trustworthiness of the mobile phone, which means that for those who used mobile media for approximately less than 5 h per day ($10^{0.7} = 5.01$), anthropomorphic language had positive effects on perceived trustworthiness. For those who used mobile media for about more than 5 h per day, there was no significant relationship between anthropomorphic language and perceived trustworthiness of the phone.

Mobile media usage interacted with the language cues in predicting users' conformity intention, $B = 0.43$, $p = .045$, 95% *CIs* [0.01, 0.86]. For those who spent less time on mobile media, anthropomorphic language evoked greater conformity intention, while for those who spent more time on mobile phones, anthropomorphic language evoked less conformity intention (Fig. 1). Mobile media usage did not interact with language or modality cues in predicting other types of social responses.

RQ2 asked how users' intensive phone use would interact with the language or the modality cues of the phone in predicting users' social responses. Moderation analyses revealed that participants' intensive phone use interacted with the modality of the phone in predicting social presence, $B = 0.95$, $p = .013$, 95% *CIs* [0.21, 1.69], perceived attraction of the phone, $B = 1.17$, $p = .019$, 95% *CIs*

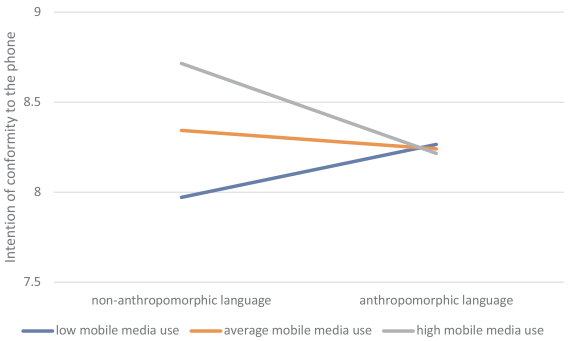

Fig. 1. The interaction between language and mobile media usage on conformity intention (RQ1). Note: Low mobile media use: 1 SD below the mean. High mobile media use: 1 SD above the mean.

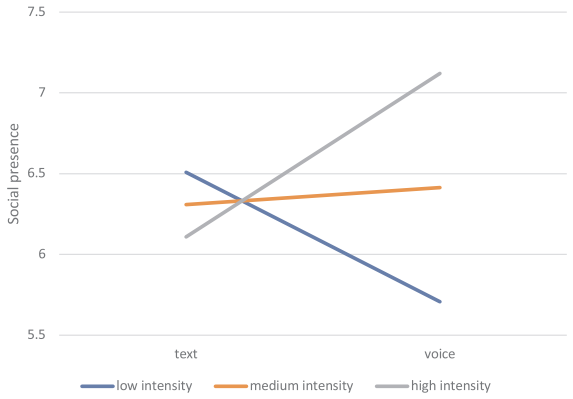

Fig. 2. Interaction between the modalities and intensive phone use on social presence (RQ2). Note: Low intensity: 1 SD below the mean. High intensity: 1 SD above the mean.

[0.20, 2.14], perceived trustworthiness of the phone, $B = 0.76$, $p = .016$, 95% *CIs* [0.15, 1.38], and users' conformity intention, $B = 0.80$, $p = .013$, 95% *CIs* [0.18, 1.42]. That is, for those who were less intensive smartphone users, text-based messages evoked greater levels of social presence, perceived attraction, perceived trustworthiness, and users' conformity intention than voice-based messages, while for those who were more intensive smartphone users, text-based messages evoked lower levels of these types of social responses than voice-based messages. The interaction effects were shown in Figs. 2–5. Intensive phone use did not interact with language or modality cues in predicting other types of social responses.

RQ3 asked how participants' gender would interact with the language or the modality cues in predicting users' social responses. Results showed that although the interaction effects were not statistically significant, females ($M = 8.77$, $SD = 1.39$) were more likely than males ($M = 7.85$, $SD = 1.68$) in their intention to conform to the mobile phone, F (1, 99) = 8.63, $p = .004$, partial $\eta^2 = 0.08$, meaning that there was main effect of gender on users' conformity intention.

5. Discussion

This study seeks to understand how the CASA paradigm can be applied to the human-phone interaction and how different language and modality cues have distinct effects on users' social responses to smartphones. This study takes individual factors into consideration and exhibits that people's mobile media use experiences play an important moderating role in users' social reactions to smartphones.

5.1. Language cues

In this study, anthropomorphic language did not differ from non-anthropomorphic language in leading to users' social responses. The results diverge from previous studies where anthropomorphic computer images had positive effects on their credibility (Gong, 2008) and anthropomorphic linguistic cues on a health website elicited more self-disclosure (Sah and Peng, 2015). One possible explanation for the non-significant differences in users' conformity intention and perceived trustworthiness was reflected in the moderation analyses, where those who spent more time on mobile media every day were less likely to show intention to conform to the mobile phone with anthropomorphic language, while those who spent less time on mobile media were more sensitive to

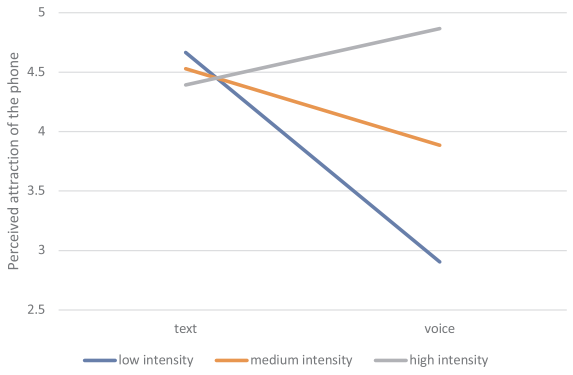

Fig. 3. Interaction between the modalities and intensive phone use on perceived attraction (RQ2). Note: Low intensity: 1 SD below the mean. High intensity: 1 SD above the mean.

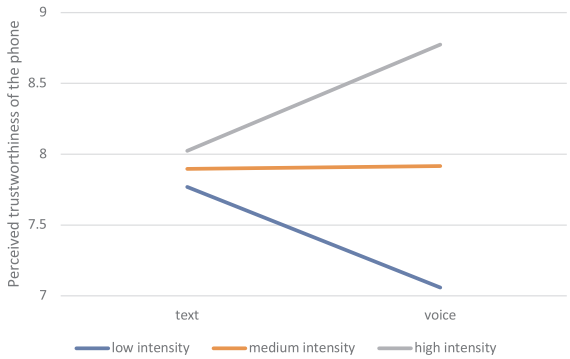

Fig. 4. Interaction between the modalities and intensive phone use on perceived trustworthiness (RQ2). Note: Low intensity: 1 SD below the mean. High intensity: 1 SD above the mean.

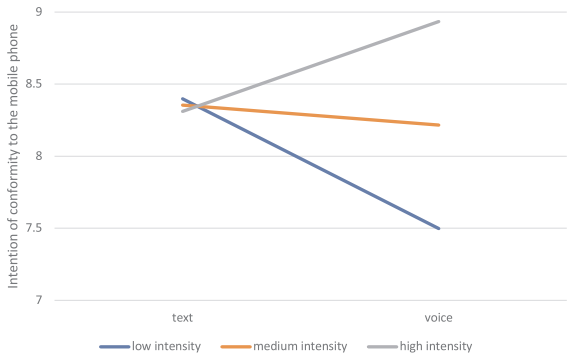

Fig. 5. Interaction between the modalities and intensive phone use on conformity intention (RQ2). Note: Low intensity: 1 SD below the mean. High intensity: 1 SD above the mean.

anthropomorphic language and were more likely to have conformity intention. In regard to users' trust in the phone, for those who spent less time on mobile media, there were positive effects of the anthropomorphic language on the perceived trustworthiness of the mobile phone, but the effects faded away for those who spent more than approximately 5 h on mobile media every day. The results imply that greater levels of mobile media usage reduced the effects of anthropomorphic linguistic cues on people's social responses to mobile technologies. The longer time people spend on mobile media, the less sensitive they are to the effects of the anthropomorphic elements.

While there were also non-significant differences in predicting social presence and perceived attraction of the phone, which is contrary to Jung and Lee's (2004) finding about the positive relationship between social cues and social attraction in HCI, it was found that users' perception of language mediated the relationship between language and these types of social responses. That is, it was indeed users' interpretation of how conversational and personal the language was that exerted stronger effects on users' social perception of the phone and social attraction of the phone. The results could complement previous work on the social influence of anthropomorphic cues by stressing the role of people's perception of language in their social interactions with media technologies.

Some may argue that examining users' perception of cues could lead scholars to focus too much attention on subjective judgments while ignoring the message features. However, scholars have responded to these concerns and discussed the necessity to use mediation models to take both media stimuli effects and psychological effects into account (see Tao and Bucy, 2007). For example, O'Keefe (2003) argued that psychological states should be assessed as potential mediators between message variations and message effects rather than as checks on messages manipulations. The current study follows these experimental designs and confirms Potter and Tomasello's (2003) finding that adding interpretation variables can increase the validity of the experimental designs, especially when the messages were not effect-labeled but effect-based (Tao and Bucy, 2007).

Apart from individual differences in mobile media usage, the task-oriented context could be another factor that explains the non-significant differences between the language cues. Given that participants were asked to react to an urgent survival scenario, the messages that were more straightforward and formal probably produced greater levels of trust and conformity intention, which might have counterbalanced the effects of anthropomorphic cues on users' reactions. Thus, future research could select a socioemotional context to examine the validity of the results.

Sundar's (2008) theorization of machine heuristics may provide another explanation for the non-significant differences. Machine heuristics suggest that if a technology presents machinelike cues on its interfaces, users may attribute objectivity and machine characteristics to the technology. Thus, it is possible that non-anthropomorphic language might have elicited some positive effects on users' trust in the mobile phone, which offset the impact of anthropomorphic language.

5.2. Modality Cues, behavioral Change, and gender differences

Voice-based and text-based modalities did not differ in users' social responses to the smartphone. The results conflict with prior research on the power of voice over text in activating users' positive attitudes toward media interfaces (Sproull et al., 1996). They are also at odds with the proposition that humans should be more sensitive to voice than to text (Nass and Brave, 2005). However, like the relationship between language cues and social responses, the effects of modality cues were affected by individual differences. Specifically, for those who used smartphones more innovatively, efficiently, and thoroughly, voice-based messages were more likely to lead them to perceive the phone as a social actor, be attracted to the phone, trust the phone, and be willing to comply with the instructions of the phone, while for those who only relied on the routine features of the phone, text-based messages were more likely to trigger these social responses. The results partially explained why voice-based messages and text-based messages did not lead to different levels of social responses. They imply that more intensive smartphone users are more comfortable receiving voice-based information, while less intensive smartphone users are more accustomed to receiving text-based information. Considering that text is still the major modality for most phone activities (e.g., reading news online, texting messages, checking emails), it is reasonable to speculate that ordinary phone users have higher dependency on text-based interactions than voice-based ones.

Neither the modality cues nor the language cues predicted users' actual behavioral change. Mobile media use experiences also did not reinforce or alleviate users' behavioral change. Actually, 84.5% of the participants conformed to all the suggestions from the smartphone. Thus, we turned to participants' responses to the open-ended questions to seek explanations. First, some participants had little knowledge about the survival scenario. They simply complied with the smartphone that provided both hints and rationale. For example, one participant wrote, "I did [accept the suggestions], because I am absolutely not a survivalist in any shape or form, so I am completely relying on the knowledge of people or databases that are better equipped to have the right answers. […] That being said, I may have been more likely to challenge what the mobile phone suggested if I had greater (or any) background knowledge in survival tactics." The second explanation for participants' conformity to the smartphone was that they had been used to trusting the information from a smartphone. One participant wrote, "For some reason I feel like when you ask a phone a question or it gives you an answer, it is usually correct." Another participant equated the information retrieval process to Google search, "I did accept all the suggestions from the mobile phone…because my generation relies a lot on Google for answers to things. I just took all the advice that it gave me because I trusted it." Overall, this study fails to reveal participants' behavioral change across different conditions. To understand participants' conformity to smartphones or other media technologies, future research can replace the current dichotomous scale with a continuous scale where participants could have more flexibility to point out the extent of their compliance.

Females were found to be more willing to conform to the mobile phone, which indicates that they were more open in accepting the suggestions from the device. The finding confirmed that there are gender differences in users' psychological responses to media technologies or media presentations (Lombard et al., 2000). It is also aligned with Lee's (2008) speculation that men were more skeptical about informational influence than women when responding to a female computer character delivering text-based messages. Despite the conjecture, it should be noted that gender did not make differences in other types of social responses. Considering that it may be risky to overgeneralize the gender differences in conformity intention only based on the current experimental setting, more research is needed to understand the role of gender in users' attitudes toward mobile phones.

5.3. Theoretical, Practical, and ethical implications

The findings of this study can help expand the scope of the CASA paradigm in that users see mobile phones as not only intermediaries, but also digital interlocutors. The results have corroborated recent studies on users' responses to smartphones as social actors (Carolus et al., 2019a; Carolus et al., 2019b; Kim, 2014). While these studies started from how users perceive specialist phones as more credible than generalist phones (Kim, 2014) and how users evaluate polite phones to be friendlier than impertinent phones (Carolus et al., 2019b), the current study compares two categories of social dimensions that have been postulated to activate users' social responses. Although no main effects were found for anthropomorphic language cues and vocal cues, the mean values of participants' responses to the measures of social presence, perceived trustworthiness of the phone, and conformity intention across the four experiment conditions were all over 5.9 on ten-point scales, indicating that participants naturally and intuitively perceived the phone as social, communicative, and credible. The findings confirmed that the scope of the CASA paradigm can be extended to human-phone interaction.

The interactions between individuals' mobile media use experiences and the cues of the smartphones support the proposition that it is not only the cues, but also the combination of individual differences and contextual factors that lead individuals to perceive media as social actors (Lombard, 2018; Xu and Lombard, 2016). The findings can complement the CASA paradigm as Nass and Moon (2000) called for more research on individual differences in users' social responses to technologies. The proposition about cues, individual differences and contextual factors can assist the theory building of the CASA paradigm and provide researchers with more hypotheses and research questions to test in the future. While much of the prior CASA research has found evidence of users' responses to technologies as social actors, future research could concentrate on when and why such responses occur (Nass and Moon, 2000), especially what the combinations are that facilitate or inhibit users' attitudinal and behavioral responses to media technologies.

The study also demonstrates that users' perception of the social cues is more important than the cues per se. While the cues can be the antecedents of users' affinity with or trust in computer interfaces (Sah and Peng, 2015), this study unfolds individuals' psychological processing of smartphones and illustrates that their interpretation of the cues has more straightforward and powerful influence on their social reactions.

Throughout the experiment, participants were asked to react to the smartphones per se. However, it remains unknown whether

they oriented their attention to the phone or other entities such as the mobile app, the voice within the app, or the app designer. Although Nass and Moon (2000) and Sundar and Nass (2000) found that it is the computer itself rather than the programmer or the networker (i.e., the person within the computer) that serves as the object of users' social attribution, the advancement of mobile technologies may have changed how users interpret the source of the interactions. One example would be that Guzman (2019) found that mobile users perceived the voices of different virtual assistants to be both in and of the machines. Thus, further research is needed to decode the target of users' social responses in human-phone interaction. One approach is to apply Solomon and Wash's (2014) theory of orientation and re-orientation to understand how technology features such as user interfaces determine users' orientation behavior.

In addition to these theoretical implications, this study can make practical contributions. For instance, to increase the perceived attraction and trustworthiness of mobile technologies, designers could customize the modality cues for both intensive mobile users and non-intensive mobile users. For those who prefer to explore the new functions of a mobile innovation, designers can consider implanting more voice-based interaction between users and the technology. Designers can also consider tailoring the language styles for users with different mobile media use habits. To increase the credibility of the messages from mobile phones, embedding more self-referential statements or exaggerated tones in the language would help those who do not spend much time on mobile devices such as elderly people learn and obtain new information.

Meanwhile, researchers should be fully aware of the ethical issues raised in these findings. For example, researchers should be cautious in manipulating the cues of the mobile phones if the information is intended for deception or discrimination. It would also be dangerous to utilize mobile phones or mobile virtual assistants when anthropomorphic language is applied to promote behavioral change like conformity. As potential solutions, designers should adopt comprehensive ethical codes and design principles that aims at clarifying the potential consequences of design, reducing social prejudices and stereotypes, and augmenting user agency in human-technology interactions (Amershi et al., 2019; Horvitz, 1999).

6. Conclusions and limitations

Technology advancements have substantially changed the way people use phones. While users may maintain social relationships with others through phones, they can also form social connection with phones per se. For example, people can easily activate mobile voice assistants by calling their names or tapping on the home button. When people use phones for video calls, they select a pleasant seating position and adjust their distance from the devices. Some people put the phones under their pillows when they go to sleep. All these life experiences manifest that we actually treat phones as if they were our companions. To fully understand our relationships with mobile phones, researchers need to explicate the psychological mechanism of our social responses to mobile phones. As was argued by Turkle (2012), humans have reached a point of seeing digital objects as both machines and creatures. Even knowing that machines should not warrant humanlike responses does not alter our reactions to them as social entities. Based on these media phenomena, this study has its value in demonstrating that the social dimensions of smartphones, along with users' perception, mobile media use experiences, and communication contexts can all contribute to users' social interactions with smartphones.

Despite the contributions, the study has its limitations. First, only 111 participants from a public university in the Northeastern U.S. participated in the experiment. The small sample size may have resulted in low statistical power of the results. To increase the external validity of the experiment, larger and more diverse samples are needed. Second, the experiment used a female voice to deliver the messages to the participants in both voice conditions. It is possible that the gender of the voice was a confounding factor in examining participants' perception and attitudes toward the smartphone. As males and females may differ in their preferences for gender-based voices, future research should balance the effects of gendered voices on users' interactions with smartphones. Third, in the voice conditions, although the participants spent most time listening to the messages via the smartphone, they had to use the buttons designed into the interfaces. These buttons had the labels such as "PREV" and "NEXT" on them, which could be considered as textual cues. Therefore, to better manipulate different conditions, scholars could consider using voice commands to avoid the influence of mixed modalities in the future.

Declaration of Competing Interest

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

References

Adkins, M., Brashers, D.E., 1995. The power of language in computer-mediated groups. Manage. Commun. Q. 8, 289–322.
Amershi, S., Weld, D., Vorvoreanu, M., Fourney, A., Nushi, B., … Teevan, J., 2019. Guidelines for human-AI interaction. In Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (p. 3). ACM.
Araujo, T., 2018. Living up to the chatbot hype: The influence of anthropomorphic design cues and communicative agency framing on conversational agent and company perceptions. Comput. Hum. Behav. 85, 183–189.
Bartneck, C., Reichenbach, J., Carpenter, J., 2008. The carrot and the stick: The role of praise and punishment in human-robot interaction. Interact. Stud. 9, 179–203.
Breazeal, C., 2003. Emotion and sociable humanoid robots. Int. J. Hum. Comput. Stud. 59 (1), 119–155.
Burgoon, J.K., Bonito, J.A., Bengtsson, B., Cederberg, C., Lundeberg, M., Allspach, L., 2000. Interactivity in human-computer interaction: A study of credibility, understanding, and influence. Comput. Hum. Behav. 16, 553–574.
Campbell, S.W., 2013. Mobile media and communication: A new field, or just a new journal. Mobile Media Commun. 1 (1), 8–13.
Carolus, A., Binder, J.F., Muench, R., Schmidt, C., Schneider, F., Buglass, S.L., 2019a. Smartphones as digital companions: Characterizing the relationship between

users and their phones. New Media Soc. 21, 914–938.
Carolus, A., Muench, R., Schmidt, C., Schneider, F., 2019b. Impertinent mobiles – Effects of politeness and impoliteness in human-smartphone interaction. Comput. Hum. Behav. 93, 290–300.
Chang, R.C.S., Lu, H.P., Yang, P., 2018. Stereotypes or golden rules? Exploring likable voice traits of social robots as active aging companions for tech-savvy baby boomers in Taiwan. Comput. Hum. Behav. 84, 194–210.
Crooks, S.M., Cheon, J., Inan, F., Ari, F., Flores, R., 2012. Modality and cueing in multimedia learning: Examining cognitive and perceptual explanations for the modality effect. Comput. Hum. Behav. 28 (3), 1063–1071.
Davis, G., 2002. Anytime/anyplace computing and the future of knowledge work. Commun. ACM 45, 67–73.
DeBaillon, L., Rockwell, P., 2005. Gender and student-status differences in cellular telephone use. Int. J. Mobile Commun. 3 (1), 82–98.
Derks, D., Bakker, A.B., 2014. Smartphone use, work-home interference, and burout: A diary study on the role of recovery. Appl. Psychol. 63, 411–440.
de Souza e Silva, A., 2013. Location-aware mobile technologies: Historical, social, and spatial approaches. Mobile Media Commun. 1 (1), 116–121.
Deyan, G., 2019 (March 28). 60+ smartphone statistics in 2019. Techjury. Retrived from https://techjury.net/stats-about/smartphone-usage/.
Dredze, M., Paul, M.J., Bergsma, S., Tran, H., 2013 (June). Carmen: A twitter geolocation system with applications to public health. In AAAI workshop on expanding the boundaries of health informatics using AI (HIAI) (pp. 20–24).
Duffy, B.R., Zawieska, K., 2012 (September). Suspension of disbelief in social robotics. In IEEE RO-MAN: The 21st IEEE International Symposium on Robot and Human Interactive Communication. Paris, France.
Fogg, B.J., 2002. Computers as persuasive social actors. In: Persuasive computers using technology to change what we think and do. Morgan Kaufman Publishers, San Francisco, CA, pp. 89–120.
Frith, J., 2014. Communicating through location: The understood meaning of the foursquare check-in. J. Comput.-Mediated Commun. 19 (4), 890–905.
Gong, L., 2008. How social is social responses to computers: The function of the degree of anthropomorphism in computer representations. Comput. Hum. Behav. 24, 1494–1509.
Gong, L., Nass, C., 2007. When a talking face computer agent is half-human and half-humanoid: Human identity and consistency preference. Hum. Commun. Res. 33, 163–193.
Guzman, A., 2019. Voices in and of the machine: Source orientation toward mobile virtual assistants. Comput. Hum. Behav. 90, 343–350.
Harmon, E., Mazmanian, M., 2013, April. Stories of the Smartphone in everyday discourse: Conflict, tension & instability. In Proceedings of the SIGCHI conference on human factors in computing systems (pp. 1051–1060). ACM.
Horvitz, E., 1999, May. Principles of mixed-initiative user interfaces. In Proceedings of the SIGCHI conference on Human Factors in Computing Systems (pp. 159–166). ACM.
Humber, 2017. Lost at Sea. Retrieved from http://www.humber.ca/centreforteachingandlearning/instructional-strategies/teaching-methods/classroom-strategies-designing-instruction/activities-and-games/lost-at-sea.html.
Humphreys, L., 2013. Mobile social media: Future challenges and opportunities. Mobile Media Commun. 1 (1), 20–25.
iOS Human Interface Guidelines, 2017, March. iOS Human Interface Guidelines. Retrieved from https://developer.apple.com/ios/human-interface-guidelines/overview/design-principles/.
Ito, M., 2005. Mobile phones, Japanese youth, and the re-placement of social contact. In: Ling, R., Pedersen, P.E. (Eds.), Mobile Communications: Re-negotiation of the social space. Springer, London, pp. 131–148.
Jethani, S., Leorke, D., 2013. Ideology, obsolescence and preservation in digital mapping and locative art. Int. Commun. Gaz. 75, 484–501.
Johnson, D., Gardner, J., Wiles, J., 2004. Experience as a moderator of the media equation: The impact of flattery and praise. Int. J. Hum. Comp. Stud. 61, 237–258.
Joosse, M., Lohse, M., Pérez, J.G., Evers, V. (Eds.), 2013. What you do is who you are: The role of task context in perceived social robot personality. IEEE, pp. 2134–2139.
Jung, Y., Lee, K.M., 2004. Effects of physical embodiment on social presence of social robots. Proc. Presence 80–87.
Jung, M., Martelaro, N., Hoster, H., Nass, C. 2014, June. Participatory materials: Having a reflective conversation with an artifact in the making. In Proceedings of the 2014 conference on Designing Interactive systems (pp. 25–34). ACM.
Katz, J.E., 2007. Mobile media and communication: Some important questions. Commun. Monogr. 74, 389–394.
Kim, K.J., 2014. Can smartphones be specialists? Effects of specialization in mobile advertising. Telematics Inform. 31, 640–647.
Kim, A., Han, J., Jung, Y., Lee, K., 2013, March. The effects of familiarity and robot gesture on user acceptance of information. In Human-Robot Interaction (HRI), 2013 8th ACM/IEEE International Conference on (pp. 159–160). IEEE.
Kim, J., Park, H.S., 2011. The effect of uniform virtual appearance on conformity intention: Social identity model of deindividuation effects and optimal distinctiveness theory. Comput. Hum. Behav. 27, 1223–1230.
Lee, E.J., 2006. When and how does depersonalization increase conformity to group norms in computer-mediated communication? Commun. Res. 33, 423–447.
Lee, E.J., 2008. Flattery may get computers somewhere, sometimes: The moderating role of output modality, computer gender, and user gender. Int. J. Hum Comput Stud. 66, 789–800.
Lee, E.J., 2010. What triggers social responses to flattering computers? Experimental tests of go and mindlessness explanations. Commun. Res. 37, 191–214.
Lee, E.J., Nass, C., 2002. Experimental tests of normative group influence and representation effects in computer-mediated communication: When interacting via computers differs from interacting with computers. Hum. Commun. Res. 28, 349–381.
Lee, K., Nass, C., 2005. Social-psychological origins of feelings of presence: Creating social presence with machine-generated voices. Media Psychol. 7, 31–45.
Lee, K.M., Peng, W., Jin, S.A., Yan, C., 2006. Can robots manifest personality? An empirical test of personality recognition, social responses, and social presence in human-robot interaction. J. Commun. 56, 754–772.
Liew, T.W., Tan, S.M., 2018. Exploring the effects of specialist versus generalist embodied virtual agents in multi-product category online store. Telematics Inform. 35 (1), 122–135.
Liu, C., 2010. Measuring and prioritizing value of mobile phone usage. Int. J. Mobile Commun. 8 (1), 41–52.
Lombard, M., 2018. Presence past and future: Reflections on 25 years of presence technology, scholarship, and community. In: Guzman, A.L. (Ed.), Human-machine communication: Rethinking communication, technology, and ourselves. Peter Lang Publishing.
Lombard, M., Ditton, T., 1997. At the heart of it all: The concept of presence. J. Comp.-Mediated Commun. 3. https://doi.org/10.1111/j.1083-6101.1997.tb00072x.
Lombard, M., Reich, R.D., Grabe, M.E., Bracken, C., Ditton, T., 2000. Presence and television. The role of screen size. Hum. Commun. Res. 26 (1), 75–98.
Martínez-Miranda, J., Pérez-Espinosa, H., Espinosa-Curiel, I., Avila-George, H., Rodríguez-Jacobo, J., 2018. Age-based differences in preferences and affective reactions towards a robot's personality during interaction. Comput. Hum. Behav. 84, 245–257.
Material Design, 2017, March. Material Design. Retrieved from https://material.io/guidelines/.
McCroskey, J., McCain, T., 1974. The measurement of interpersonal attraction. Speech Monogr. 41, 261–266.
Mihailidis, P., 2014. A tethered generation: Exploring the role of mobile phones in the daily life of young people. Mobile Media Commun. 2 (1), 58–72.
Misra, S., Cheng, L., Genevie, J., Yuan, M., 2016. The iPhone effect: The quality of in-person social interactions in the presence of mobile devices. Environ. Behav. 48, 275–298.
Nardi, B.A., Whittaker, S., Bradner, E., 2000, December. Interaction and outeraction: Instant messaging in action. In Proceedings of the 2000 ACM conference on Computer supported cooperative work (pp. 79–88). ACM.
Nass, C., 2004. Etiquette quality: Exhibitions and expectations of computer politeness. Commun. ACM 47, 35–37.
Nass, C., Brave, S., 2005. Wired for speech: How voice activates and advances the human-computer relationship. MIT Press, Cambridge, MA.
Nass, C., Fogg, B.J., Moon, Y., 1996. Can computers be teammates? Int. J. Hum. Comp. Stud. 45, 669–678.
Nass, C., Lombard, M., Henriksen, L., Steuer, J., 1995a. Anthropocentrism and computers. Behav. Inform. Technol. 14, 229–238.
Nass, C., Moon, Y., 2000. Machines and mindlessness: Social responses to computers. J. Social Issues 56, 81–103.
Nass, C., Moon, Y., Fogg, B.J., Reeves, B., Dryer, C., 1995b, May. Can computer personalities be human personalities? In Conference companion on Human factors in

computing systems (pp. 228–229). ACM.
Nass, C., Steuer, J., 1993. Voices, boxes, and sources of messages: Computers and social actors. Hum. Commun. Res. 19, 504–527.
Nass, C., Steuer, J., Tauber, E.R., 1994. Computers are social actors. Hum. Factors Comput. Syst. 94, 72–78.
National Institute of Health, 2017, March. What is voice? What is speech? What is language? Retrieved from https://www.nidcd.nih.gov/health/what-is-voice-speech-language.
Okazaki, S., Hirose, M., 2009. Does gender affect media choice in travel information search? On the use of mobile Internet. Tourism Manage. 30, 794–804.
O'Keefe, D.J., 2003. Message properties, mediating states, and manipulation checks: Claims, evidence, and data analysis in experimental persuasive message effects research. Commun. Theory 13, 251–274.
Oh, J., Robinson, H.R., Lee, J.Y., 2013. Page flipping vs. clicking: The impact of naturally mapped interaction technique on user learning and attitudes. Comput. Hum. Behav. 29, 1334–1341.
Panda, A., Jain, N.K., 2018. Compulsive smartphone usage and users' ill-being among young Indians: Does personality matter? Telematics Inform. 35, 1355–1372.
Pew Research Center, 2018. Mobile fact sheet. Pew Research Center: Internet and Technology. Retrieved from http://www.pewinternet.org/fact-sheet/mobile/.
Potter, W.J., Tomasello, T.K., 2003. Building upon the experimental design in media violence research: The importance of including receiver interpretations. J. Commun. 53, 315–329.
Przybylski, A.K., Weinstein, N., 2013. Can you connect with me now? How the presence of mobile communication technology influences face-to-face conversation quality. J. Social Pers. Relat. 30, 237–246.
Reeves, B., Nass, C., 1996. The media equation: How people treat computers, television, and new media like real people and places. CSLI Publications, CA, U.S.
Sah, Y.J., Peng, W., 2015. Effects of visual and linguistic anthropomorphic cues on social perception, self-awareness, and information disclosure in a health website. Comput. Hum. Behav. 45, 392–401.
Schafer, J.L., 1999. Multiple imputation: A primer. Stat. Methods Med. Res. 8 (1), 3–15.
Schlosser, F.K., 2002. So, how do people really use their mobile devices? An interactive study of wireless technology use. J. Organ. Behav. 23, 401–423.
Scolari, C.A., Aguado, J.M., Feijoo, C., 2012. Mobile media: Towards a definition and taxonomy of contents and applications. Int. J. Interact. Mobile Technol. 6, 29–38.
Short, J., Williams, E., Christie, B., 1976. The social psychology of telecommunications. Wiley, London.
Sproull, L., Subramani, M., Kiesler, S., Walker, J.H., Waters, K., 1996. When the interface is a face. Hum. Comput. Interact. 11, 97–124.
Statista, 2019, August 21. Number of mobile phone users worldwide from 2015-2020. Statista. Retrieved from https://www.statista.com/statistics/274774/forecast-of-mobile-phone-users-worldwide/.
Sundar, S., 2008. The MAIN model: A heuristic approach to understanding technology effects on credibility. In: Metzger, M.J., Flanagin, A.J. (Eds.), Digital media, youth, and credibility. The MIT Press, Cambridge, MA, pp. 73–100.
Sundar, S.S., Jia, H., Waddell, F., Huang, Y., 2015. Toward a theory of interactive media effects (TIME): Four models for explaining how interface features affect user psychology. In: Sundar, S.S. (Ed.), The handbook of the psychology of communication technology. John Wiley & Sons Inc, UK, pp. 47–86.
Sutko, D.M., de Souza e Silva, A, 2010. Location-aware mobile media and urban sociability. New Media Soc. 13, 807–823.
Takayama, L., Groom, V., Nass, C., 2009, April. I'm sorry, Dave: I'm afraid i won't do that: social aspects of human-agent conflict. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (pp. 2099–2108). ACM.
Tao, C., Bucy, E.P., 2007. Conceptualizing media stimuli in experimental research: Psychological versus attribute-based definitions. Hum. Commun. Res. 33, 397–426.
Turkle, S., 2012. Alone together: Why we expect more from technology and less from each other. Basic books, New York, NY.
Vinciarelli, A., Pantic, M., Bourland, H., 2009. Social signal processing: Survey of an emerging domain. Image Vis. Comput. 27, 1743–1759.
Wei, R., 2008. Motivations for using the mobile phones for mass communications and entertainment. Telematics Inform. 25, 36–46.
Wei, R., 2013. Mobile media: Coming of age with a big splash. Mobile Media Commun. 1 (1), 50–56.
Woods, S., Dautenhahn, K., Kaouri, C., te Boekhorst, R., Koay, K.L., Walters, M.L., 2007. Are robots like people? Relationships between participant and robot personality traits in human–robot interaction studies. Interact. Stud. 8, 281–305.
Xu, K., 2019. First encounter with robot Alpha: How individual differences interact with vocal and kinetic cues in users' social responses. New Media Soc. 21, 2522–2547.
Xu, K., Lombard, M., 2016. Media are social actors: Expanding the CASA paradigm in the 21st century. Paper presented in the 2016 Annual Conference of the International Communication Association.
Zhao, S., 2003. Toward a taxonomy of copresence. Presence Teleoperators Virtual Environ. 12, 445–455.
Zhong, B., 2013. From smartphones to iPad: Power users' disposition toward mobile media devices. Comput. Hum. Behav. 29, 1742–1748.